STAFFORDSHIRE UNIVERSITY

BUSINESS SCHOOL

ECONOMICS

THE RELATIONSHIP BETWEEN OUTPUT GAP AND EXCESS LIQUIDITY IN EUROPEAN TRANSITION ECONOMIES

ALBULENË KASTRATI

Ph.D. 2015

Abstract

The aim of this thesis is to investigate the definition and measurement of the potential output and the output gap, to examine the definition and determinants of the size and the behaviour of excess liquidity in the banking sector in transition economies and last, to explore the theoretical and empirical links between the output gap and excess liquidity. Estimating potential output and the corresponding output gap is deemed as particularly relevant for the conduct of monetary and fiscal policy as well as in providing useful insights regarding the state of the economy, especially for periods with below potential production. One of the most important factors associated with a large output gap may be the efficiency of financial systems in transition economies, in particular the behaviour of banks, as they represent the main source of finance with capital markets being in their infancy. Access to banking loans is one of the main sources of finance for a firm's growth, while excess liquidity holdings are a feature of several European transition economies. Thus, another research question addressed in this thesis is why do banks in transition economies continuously accumulate excess reserves in the face of seemingly profitable loans to invest in, relative to developed countries, which are also costly to maintain. Criticizing the prior definitions of excess liquidity in the existing literature as imprecise, another theoretical added value of this thesis consists on clarifying the excess liquidity concept and subsequently redefining its measurement. Augmenting the model with risk, regulatory and institutional variables, the findings suggest that involuntary factors that are mainly outside of banks' control (e.g. lack of credit demand, poor institutional framework) prevail in inducing banks to accumulate excess reserves, relative to the precautionary motive. Treating potential endogeneity issue between excess liquidity and the output gap, the last finding of this research proclaims that rather being in a causal relationship, both are actually correlated via the common observed and unobserved determinants. Due to the lack of a clear transmission mechanism from one to the other factor, rather than pushing banks to reduce excess liquidity to extend lending and consequently to reduce the output gap, a policy implication of the thesis is that other factors outside of the system that are capable of pushing both variables in the desired direction need to be identified. The avoidance of a policy where banks are pushed to reduce excess liquidity is the main policy implication of this chapter, since that may lead into a new wave of bad loans rather than impetus to economic activity.

Contents

List of tables

List of figures

Abbreviations

ACF – Autocorrelation Function
ADF – Augmented Dickey-Fuller
AMADEUS – Analog Modelling and Design Using a Symbolic Environment
AO – Actual Output
AR – Autoregressive
ARIMA – Autoregressive Integrated Moving Average
BEEPS – Business Environment and Enterprise Performance Survey
BN – Beveridge and Nelson
CAR – Capital Adequacy Ratio
CBK – Central Bank of the Republic of Kosovo
CFR – Common Factor Restrictions
CIS – Commonwealth of Independent States
CPI – Consumer Price Index
DF-GLS – Dickey-Fuller Generalised Least Squares
EBRD – European Bank for Reconstruction and Development
EC – European Commission
ECB – European Central Bank
EL1 – Excess Liquidity (above reserve required ratio)
EL2 – Excess Liquidity (above mandatory liquidity ratio)
EU – European Union
FDI – Foreign Direct Investment
FE – Fixed Effects
FGLS – Feasible Generalised Least Squares
FP – Food Prices
FYR – Former Yugoslavian Republic
GDP – Gross Domestic Product
GFC – Global Financial Crisis
GLS – Generalised Least Squares
GMM – Generalised Method of Moments
GNP – Gross National Product
HP – Hodrick-Prescott
IID – Individually and Independently Distributed
IMF – International Monetary Fund
LA – Liquid Assets
LCR – Liquid Assets to Net Cash Outflows
LRI – Legal Rights Index
MA – Moving Average
MLE – Maximum Likelihood Estimator
NAIRU – Non-Accelerating Inflation Rate of Unemployment
NPL – Non-Performing Loans
NSF – Net Stable Funding
OECD – Organisation for Economic Co-operation and Development

OG –Output Gap
OLS – Ordinary Least Squares
OP – Oil Prices
PACF – Partial Autocorrelation Function
PCBH – ProCredit Bank Holding
PCSE – Panel Corrected Standard Errors
PO – Potential Output
PP – Phillips-Perron
RBI – Raiffesien Bank International
RL – Required (mandatory) Liquidity
RLE – Rule of Law Estimate
RRR – Reserve Required Ratio
RWA – Risk Weighted Assets
SAK – Statistical Agency of Kosovo
SEE – South-Eastern Europe
SUR – Seemingly Unrelated Regressions
SURECM – Seemingly Unrelated Regressions Error Correction Model
SVAR – Structural Vector Autoregression
TFP – Total Factor Productivity
UC – Unobserved Components
ULC – Unit Labour Cost
VAT – Value Added Tax
VOL – Volatility
WEO – World Economic Outlook

Acknowledgements

My deepest gratitude goes to my honoured professors, Nick Adnett, Geoff Pugh and Valentin Toçi, whose commitment, valuable suggestions and critiques have respected my PhD thesis beyond its merit. I have gained much more than professorial guidance from them. They were role models of nobility and integrity and I will always look up to their example in my future academic career.

It goes without saying that the indispensable contribution of Professor Iraj Hashi and Economics Department in Staffordshire University and Open Society Foundation has the merit for engendering qualified educational staff in Kosovo for more than a decade.

I am deeply grateful to those who offered me collegial guidance and support over different stages of my thesis, such as my sister Urtinë Kastrati and my cousin Frashër Loshaj, my fellow scholars Biljana Petkovska and Jane Bogoev, my colleague and tutor Adam Gersl, and Albulena Thaçi for technical assistance. I would like to also thank two very dear persons to me, Jenny and Marion, whose generosity and willingness to help brought much more than administrative contribution to this programme and made us feel closer to home.

I would like to also express my appreciation to the kind people of Stoke-on-Trent, Yoga club and Volleyball team in Staffordshire University who made this experience much more enjoyable.

Finally and most importantly, I am forever grateful to my parents who paved the path for me, taught me the importance of education and always believed in me, which means so much to a child.

Chapter 1. Research context

The aim of this thesis is to investigate the relationship between the output gap and excess liquidity in European transition economies. The main aims of this research project are the following. The first aim of the thesis consists in investigating potential output and the output gap in European transition economies, with special reference to the Czech Republic, Estonia and Kosovo. The second aim is to examine the definition and determinants of the size and behaviour of excess liquidity in the banking sector in these economies. Finally, the last aim of the thesis is to explore the theoretical and empirical links between the output gap and excess liquidity in these economies. Based upon the findings generated, the ultimate goal of this research is to develop and disseminate policy recommendations related to these questions for European transition countries.

A considerable amount of research has already investigated macroeconomic developments, financial stability, banking efficiency and credit constraints in transition economies. However, little research has been directed at investigating the extent of underutilised resources in European transition countries in terms of estimating the size of the output gap, similarly only a few studies have investigated the excess liquidity phenomenon in transition context, and no other research has linked output gap and excess liquidity together. These issues, initially presented in this chapter, will be further investigated in this thesis.

1.1. Introduction

The aim of this chapter is to present the context of the thesis, explain its structure and establish the originality of the research programme undertaken. The evidence built in this chapter on the presence of output gaps and excess liquidity in European transition economies raise research questions that are novel in the current literature and hence contribute to closing gaps in current knowledge. The rest of the chapter is organized as following. The second section presents a summary of the context of the thesis. The third section presents evidence for particular macroeconomic characteristics that serve as evidence on the presence of output gaps in European transition economies. Section 1.4 will be particularly concerned on building the evidence on the presence and persistence of the underutilised financial resources, i.e. excess liquidity. The fifth section summarises specific banking sector characteristics dealing with banking risks and costs, followed by a section discussing the obstacles to accessing credit in European transition economies. Section 1.7 will elaborate on the reasons for selecting three particular European transition countries to represent the region. Based on the gaps identified and evidence presented, the eighth section will present the main research questions of the thesis, whereas how these are addressed in the thesis will be explained in in the ninth section. The chapter concludes with an explanation of the organisation of the remainder of this thesis.

1.2. The focus of the research programme

European transition economies have developed unevenly since 1989 with the more slowly developing countries (Berglof and Bolton, 2002), such as the South-Eastern European (SEE) countries, experiencing persistently high rates of unemployment (with many of the countries facing unemployment rates of around 20% or higher), chronic current account deficits (frequently over 10 percent of the GDP) and relatively low levels of bank loans relative to GDP (around 50% of GDP in European transition economies, compared to around 130% of GDP in Eurozone, table 1.7.1. and figure 1.7.1. in Appendix 1), (Deutsche Bank, 2011; ECB; 2015 European Commission, 2010). Even though most of these economies fairly rapidly attained macroeconomic and financial stability, this process has usually been followed by market and institutional failures such as weak contract enforcement, poor financial reporting, low rates of paying taxes etc., which have been identified as causing slower rates of economic growth (World Bank, 2015). Thus, assuming that there are relatively high unexploited resources in many European transition economies and based on the sluggish economic growth, together with these market deficiencies, transition economies may be considered to be operating well below their potential output, i.e. there is a large output gap (Jazbec and Kastrati, 2011). Investigating the magnitude and the factors causing the output gap in transition economies is particularly important, since under-utilized capacities are likely to be constraining economic growth.

Access to banking loans is one of the main sources of finance for a firm's growth. Calculations from the 2005 BEEPS data indicate that amongst all firms that have applied for a loan in the SEE region, 6.5% were rejected, which is not particularly high. Nevertheless, Hashi and Toci (2010) find evidence that, from all the firms in the sample, 31% may be considered as discouraged borrowers, which together with the ones who were denied credit constitute a large group of credit-rationed firms in transition economies. Using a survey of 80,000 firms, Beck and Demirguc-Kunt (2006, p. 9) find evidence that "firms in countries with higher levels of financial intermediary development, more liquid stock markets, more efficient legal systems and higher GDP per capita report lower financing obstacles". Using AMADEUS data, Konings et al. (2003) find that a significant number of firms in Poland and the Czech Republic seem to be credit constrained, "a result often observed in western market economies as well"; whereas in the less advanced transition economies like Bulgaria and Romania, credit constraints seem less present. Therefore, credit rationing, poor market and institutional performance (Bonin, 2001; Konings et al., 2003; Roaf et al., 2014), together with other deficiencies in European transition economies (e.g. their large informal economies) may have led to a continuous accumulation of liquidity above the required level, i.e. excess liquidity. Between 2000 and 2014 banks in transition economies accumulated liquid assets around 19% above the level required by the regulators (Albania around 37%, Romania 42.1%, Serbia 25% or Czech Republic around 14% of excess liquidity). Based on the respective central banks' regulations, in most cases excessive reserves are non-remunerative and even in the cases where they are remunerated the interest rates earned are very low. Holding excess funds available is costly for banks as they will be paying interest rates on deposits they have received. Moreover, nominal loan interest rates in most of the European transition economies during 2000' were above 11% (World Bank, 2015) compared to around 4.5% in the Eurozone (ECB, 2015), suggesting that loans are a relatively high-return investment for banks in transition economies. Thus, the question that arises is why do profit-maximising banks in transition economies not use all their available funds to invest in domestic loans? Moreover, accumulated excess liquidity is likely to be costly for the economy, as it means less credit extended to the private sector, slowing growth and widening the output gap. The presence of excess banking sector liquidity and financial constraints in transition economies is a signal of market imperfections, which may have two main effects. Firstly, excess liquidity may create limitations on working capital, which would prevent firms operating at full capacity. Secondly, excess liquidity may constrain investment which impedes growth and restricts the pace of capital accumulation, and, hence, the long-run rate of growth of actual and potential output.

A considerable part of the literature has investigated the general economic development of transition economies during the last twenty years (Berglof and Bolton, 2002; Cottareli et al., 2005; Roaf et al., 2014). Many other studies, have investigated the level of financial development in transition economies (Bonin, 2001; Claessens et al., 2001; Reininger, 2002; Caviglia et al., 2002). Furthermore, a vast part of the literature was also concerned with credit rationing theory and financial constraints in transition economies (Bratkowski et al., 2000; Konings et al., 2002; Beck et al., 2006; Inderst and Mueller, 2008), which may partly explain the relatively low lending levels, however, only a few studies have investigated the excess liquidity phenomenon in transition context (Agenor et al., 2004; Saxegaard, 2006; Khemrraj, 2007; Primus, 2013) and none examine its continuous presence in the region of European transition economies. On the other side, several studies have estimated potential output and the output gap in transition economies (Room, 2001; Zaman, 2002; Darvas and Vadas, 2003; Benes and Ndiaye, 2004; Antonicova and Hucek, 2005; Tsalinski, 2007), however, to the best of our knowledge, no other study has investigated the links between output gap and excess liquidity. Therefore, the research reported in this thesis addresses this knowledge gap.

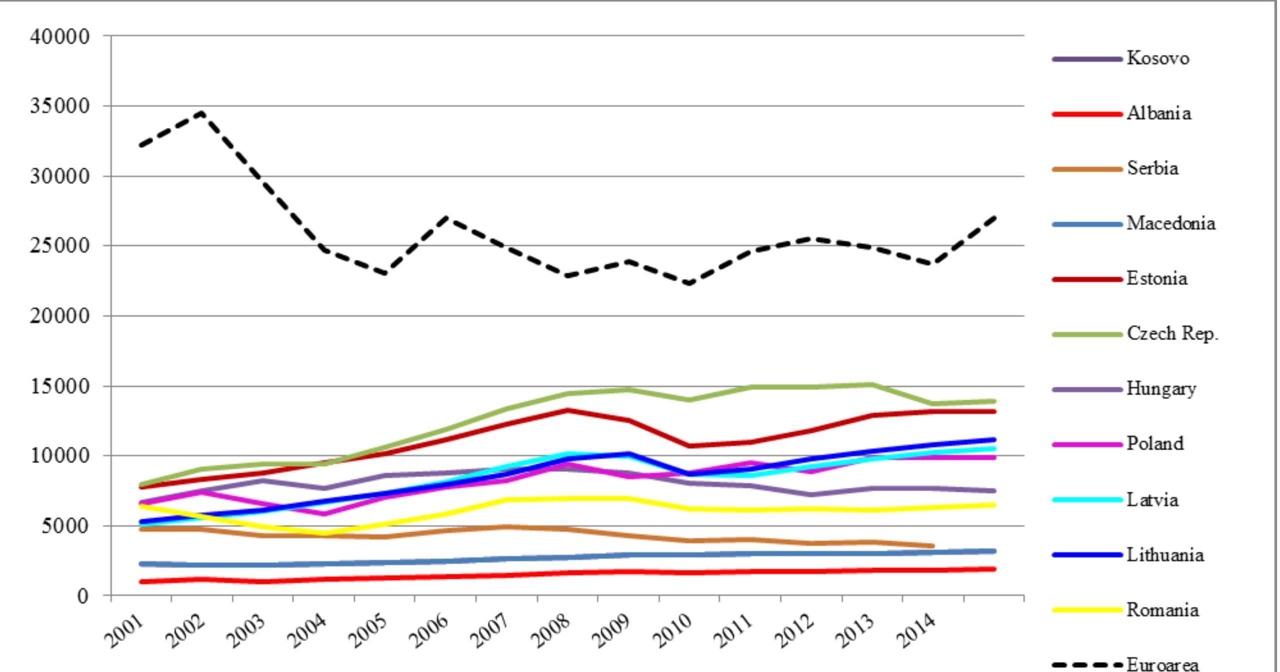
The next section will elaborate to what extent European transition economies are operating with underutilised resources and large output gaps. The following section will address the question of whether banks in European transition economies issue less credit, other things being equal, than banks in advanced European economies. Section 1.5 will investigate if the funding costs of banks in European transition countries are: higher; whether they suffer from a shortage of funds to invest; are their returns lower; is the risk perceived higher or does the problem lie in market and institutional deficiencies?

1.3. Evidence concerning the size of the output gaps in European transition economies

Most of the European transition economies have been successful in implementing reforms and enhancing the macroeconomic framework and have opened up their financial sectors, thus paving the way for economic revival (Falcetti et al., 2005). Some of these countries also representing early reformers in the transition process, like the Czech Republic, Hungary or Poland represent the most developed economies with the largest size of the real GDP per capita (Roaf et al., 2014), whereas other European transition economies which were relatively late comers to the transition process like Albania or Kosovo represent countries in this region with the lowest GDP per capita[1] and economic development (figure 1.1; see table A1.1 in Appendix 1 for numerical data).

[1] The GDP per capita in constant prices is taken from the IMF WEO (2015) database, where all the data were expressed in national currency. The real GDP per capita in national currency (non-euro countries) was multiplied by the exchange rate of the last day of December of respective years. The same procedure was followed for Euroarea data, which data were taken from Tradingeconomics (2015) in U.S. Dollars then converted into Euros.

Figure 1.1. Real GDP per capita



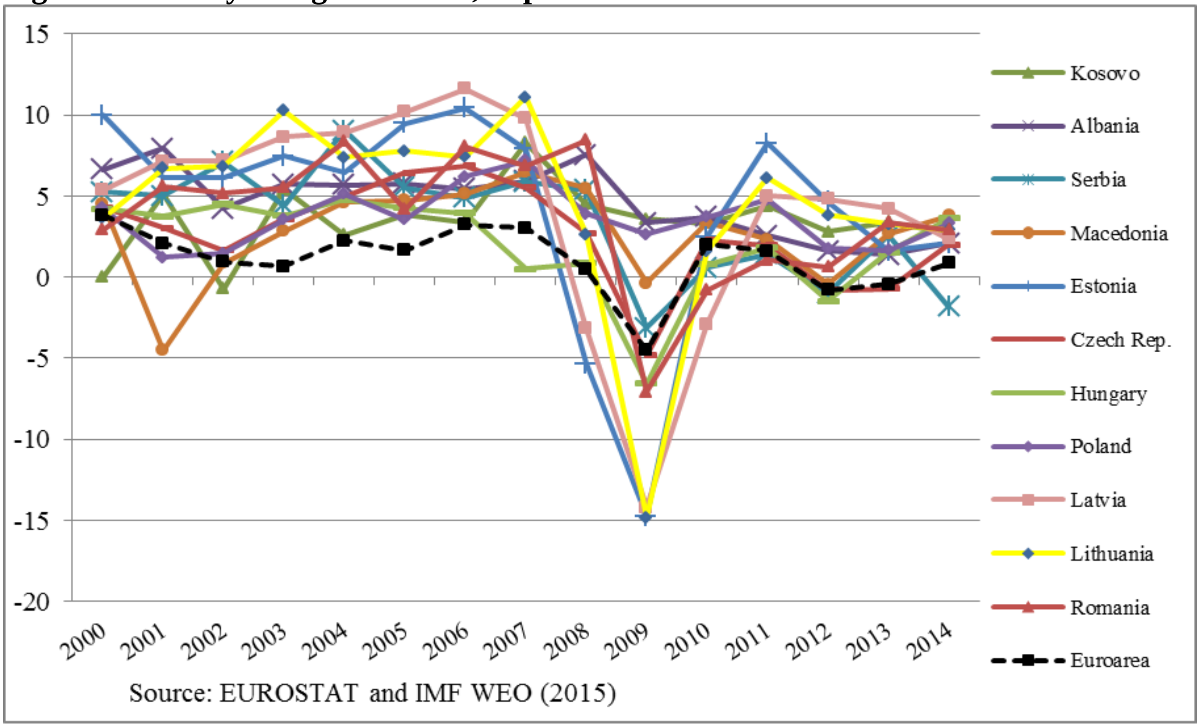
Source: IMF WEO and Tradingeconomics (2015)

Many of the European transition countries achieved a period of robust growth from the early 2000s up to 2007, experiencing the effects of the Global Financial Crisis (GFC) with recession or slower growth during 2008 and 2009, and a recuperating period afterwards (figure 1.2; see Appendix 1 for the numerical data). During the observed period (2000 – 2014), the selected countries in figure 1.2., grew at an average real growth rate of 3.5 percent, with the highest average of 4.6 percent marked in Albania and the lowest in Hungary of 2.0 percent. The European transition economies average growth rate appears as well above the average of the Euroarea region of 1.1 percent.[2] In the recent years, the European transition economies' potential growth rate may have increased due to the combined effects of the structural reforms, improved macroeconomic stability and the EU legislative harmonisation. These processes started in the mid-1990s for the more developed European transition economies and five or ten years later in the late comers to the transition process. Nevertheless, given the relatively low base of GDP, the growth of the European transition economies seems

[2] The Euroarea economic indicators presented in figures and tables of chapter 1 represent the advanced European economies and serve as a benchmark for European transition economies. The European transition economies include: Albania, Bulgaria, Czech Republic, Hungary, Poland, Romania, Slovakia; the Baltic countries, namely Estonia, Lithuania and Latvia; and the successor states of the former Yugoslavia, namely Bosnia and Herzegovina, Croatia, Kosovo, FYR Macedonia, Montenegro, Serbia and Slovenia (IMF, 2000). The Commonwealth of Independent States (CIS) countries are not included in the thesis analysis.

to be relatively sluggish in terms of catching up with the advanced economies in Europe (see table A1.2. in Appendix 1 for numerical data).

Figure 1.2.Yearly real growth rate, in percent

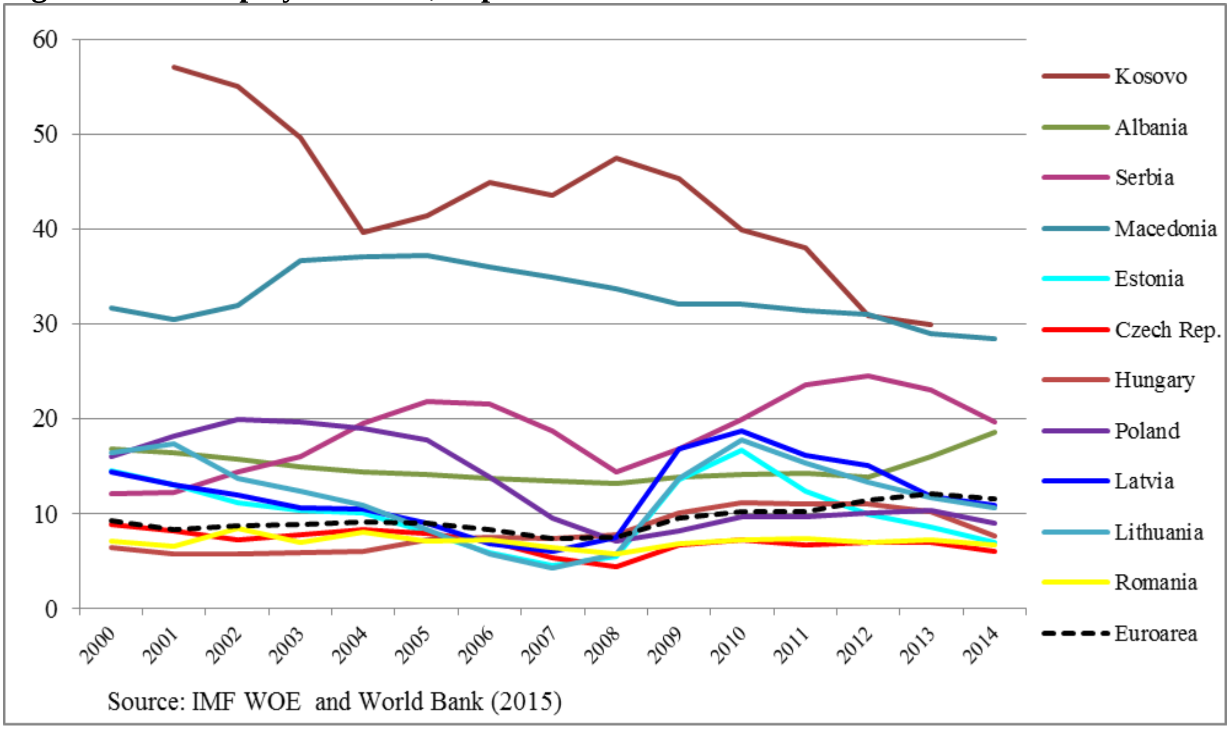


Source: EUROSTAT and IMF WEO (2015)

The question why these countries are generally lagging behind the more developed countries of Europe and how large is their output gap has not been answered to date. Therefore, the initial issues raised in this thesis will be: how large is the output gap and what are its determinants. Given that the potential output in economies is not directly observable and difficult to measure, and hence the output gap is also unobservable, policymakers need other economic indicators to provide an accurate reading of the overall demand pressure in the economy. Among the indicators usually used as proxies of the output gap are unemployment rate, capacity utilization, labour shortages, average hours worked and average hourly earnings, money and credit growth, and inflation relative to expectations (Shepherd and Driver, 2003; Jahan and Mahmud, 2013). However, not all these indicators may be good proxies in the context of economic transition. For example, given the relatively high level of structural unemployment and the rapid structural changes in transition economies, labour shortages may not be a good proxy for the output gap. Similarly, the inflation relative to expectations indicator may also not be an appropriate output gap proxy, given that the inflation rate in many of these countries mainly reflects the behaviour of import prices. Further, because of the secular rise in credit growth and the underdeveloped nature of the financial intermediation in European transition economies, money and credit growth may also not represent an appropriate output gap proxy. Moreover, data on average hours worked and average hourly earnings are not readily available for most of the European transition economies. Instead, in order to initially assess whether all the growth potential has been fulfilled in this region, other proxies that are arguably more appropriate for European transition economies, such as unemployment rate, capacity utilisation, current account balance and the credit to GDP ratio will be presented in the following paragraphs.

One of the most important indicators of the size of the output gap is the unemployment rate of a country. As can be seen in figure 1.3 below, the unemployment rate in the selected transition countries is relatively high. The Czech Republic and Hungary rank as the countries with the lowest average unemployment rate (2000 – 2014) of 7.1 and 8.1 percent, respectively, followed by a 'medium' unemployment rate in Serbia and Poland of 18.6 and 13.2 percent respectively. Countries like Kosovo and Macedonia with an average unemployment rate of 43.3 and 32.9 respectively, lead the list of the countries with the highest unemployment rate within the transition region as well as in Europe. Overall, the average unemployment rate of the selected European transition in table 1.3. for the last 14 years is 17.8 percent, which is well above the same average in Euroarea of around 9 percent (see Appendix 1 for the table of data). The relatively high rate of unemployment in European transition economies represents an indicator that one of the most important factors of production seems to be hugely under-utilized in these countries. Furthermore, the persistence of the high levels of unemployment through the years also indicates that the job growth has been relatively slow to accommodate the net inflow of workers into the labour force.

Figure 1.3. Unemployment rate, in percent



The output gap represents the state of the economy at the aggregate level, whereas the capacity utilization rate of the manufacturing firms is one of the most important proxies of output gap at firm level (Graff and Sturm, 2010).[3] Moreover, capacity utilization is a particularly relevant factor in these countries, considering that many of the European transition economies (especially SEE countries) face a relatively high trade deficit and lack of competitiveness (Havolli and Kastrati, 2012). Descriptive statistics presented in table 1.1 suggest that, on average, around 20 percent of the manufacturing capacity of firms in the selected European transition economies remains unutilized, indicating the presence of a considerable output gap at firm level.

[3]"Capacity utilization is defined as the ratio of actual output to some measure of potential output given a firm's short-run stock of capital and other fixed inputs in the short run" (Nelson, 1989, p.1). "Capacity utilization captures the output gap between actual output and capacity output" (Greboval and Munro, 1999, p. 1). "The capacity utilization rate that can be inferred from these surveys is an important business cycle indicator, as it relates directly to the pressure on current capacity to produce goods and services" (Graff and Sturm, 2010, p. 10).

Table 1.1. The average capacity utilization rate of firms

Country/ Year	2002	2005	2007	2009	Average*
Albania	82.0	76.5	72.0	76.3	*76.7*
Poland	81.4	83.6	…	83.1	*82.7*
Romania	85.3	86.8	…	81.1	*84.4*
Serbia	78.2	76.7	…	68.3	*74.4*
Moldova	72.3	77.4	…	57.3	*69.0*
B&H	79.5	81	…	74.2	*78.2*
MKD	81.1	80.6	…	74.8	*78.8*
Estonia	77.7	82.2	…	79.4	*79.8*
Czech	88.5	87.4	…	88.7	*88.2*
Hungary	82.9	80.2	…	79.8	*81.0*
Latvia	79.2	79.4	…	73.6	*77.4*
Lithuania	81.4	83.9	…	78.9	*81.4*
Slovakia	86.6	88.2	…	85.7	*86.8*
Slovenia	81.7	85.7	…	84.6	*84.0*
Bulgaria	79.8	84.6	76.0	83.8	*81.1*
Croatia	82.9	83.2	78.2	77.7	*80.5*
Monenegro	97.0	76.1	…	75.9	*83.0*
Average	*82.2*	*82.0*	*75.4*	*77.8*	*80.4*

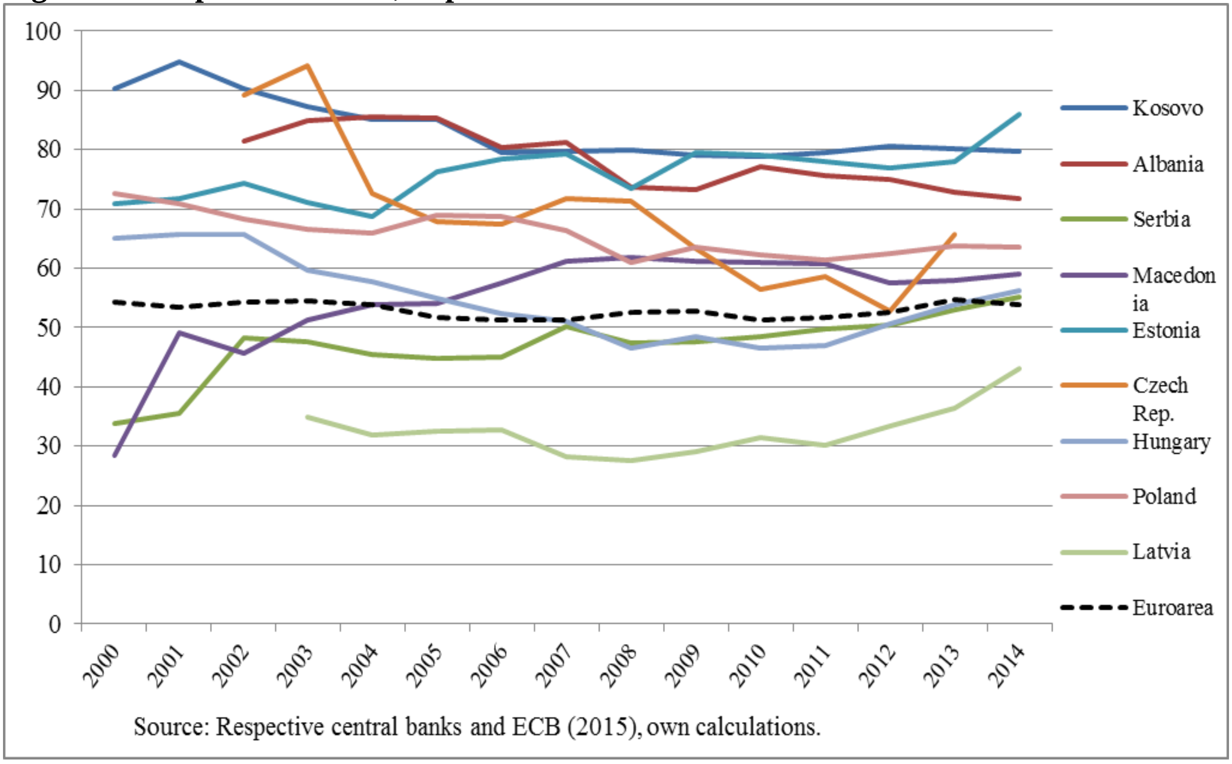
*Note: The calculated average of all the responding firms in respective years excludes firms with no answer (in BEEPS questionnaire coded as '-9').

Source: BEEPS Panel Questionnaire (2009)

As depicted in figure 1.5, an advantageous feature of banks in transitions countries is that their external risk is lessened as bank lending is usually financed by domestic deposits, which are also relatively stable funds compared to other sources like capital markets abroad (see table 1.5 in appendix 1 for the numerical data). [4] The average deposits to assets ratio (2000 – 2014) in the Euroarea was 52.9 percent (63 percent in European transition economies), meaning that the advanced economies of Europe utilise less domestic deposits and rely more on debt and other derivative instruments, which are more volatile and riskier funds.

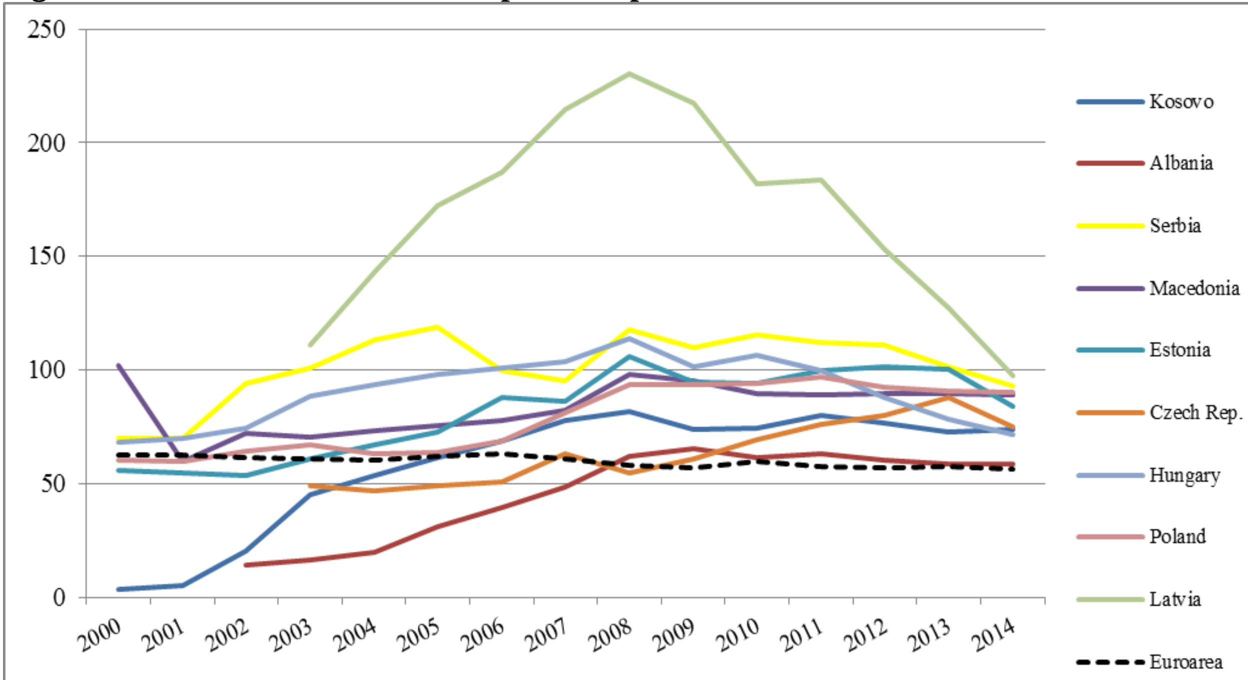
[4] The banking system in these countries was relatively stable, even after the GFC turmoil, with a few exceptions in Baltic countries. One of the reasons that the banking system in European transition economies survived the GFC relatively easier that the banks in advanced economies in Europe is the very fact that these banks were less integrated into global markets thus less exposed to the external and contagion risk.

Figure 1.4. Deposits to assets, in percent



Furthermore, the availability of domestic funds as well as the lack of capital markets enabled banks in European transition economies to maintain a simplistic business model: taking domestic deposits from one side and issuing loans on the other side. As can be noted in figure 1.6, while banking deposits comprise the largest share of assets (or liabilities), private sector loans comprise around 60 percent of deposits, meaning that loans are mainly financed by deposits, leaving around 40 percent of deposits for other investments. Amongst the countries with the lowest share of loans to deposits (on average) are Kosovo and Albania. The fairly low ratio especially at the beginning of 2000s mainly reflects the low level of banking sector development. In countries like Serbia, Latvia or even Hungary in some years the credit to deposits ratio exceeds 100 percent, meaning that they have utilised other funds in addition to deposits to satisfy the demand for loans.

Figure 1.5. Private sector loans to deposits, in percent



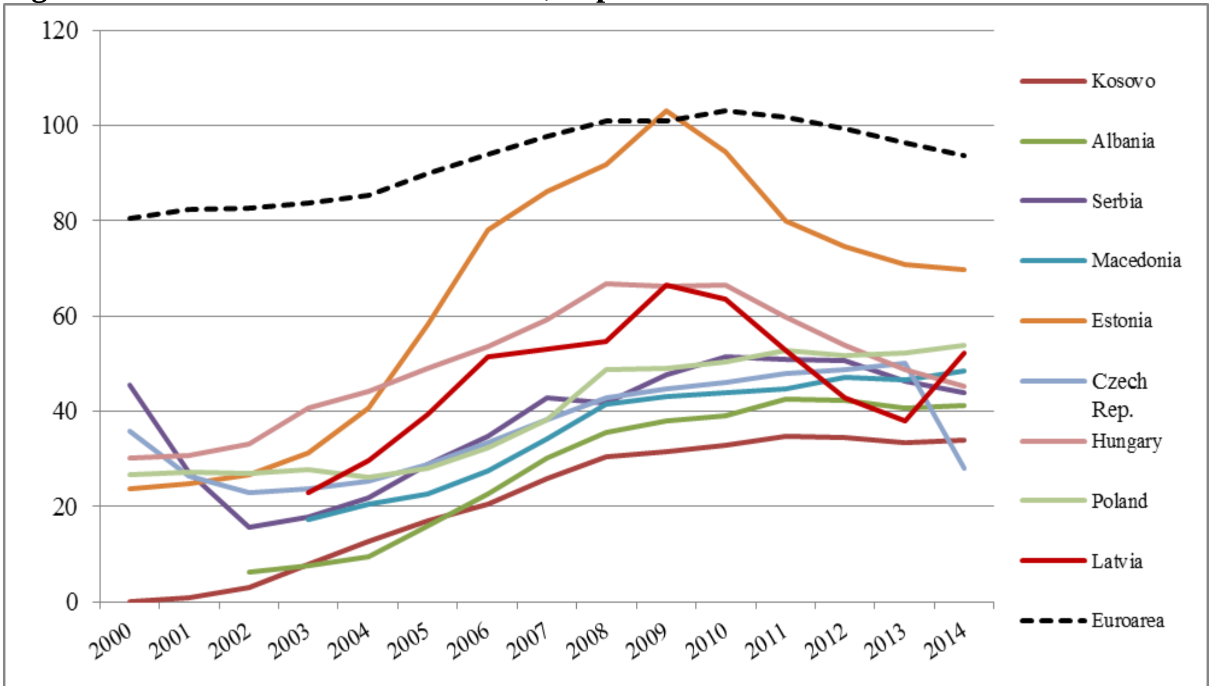
Source: Respective central banks and ECB (2015), own calculations.

Despite the relatively high ratio of deposits to assets when compared to that in the Euroarea economies (see figure 1.5), even after ten years, financial intermediation in many European transition economies does not appear to be equally strong. Even though from the mid-2000s to 2014 the degree of financial intermediation (loans and deposits) increased significantly, the average rate of private sector credit to GDP did not exceed 45 percent in these countries. This level of credit intermediation may be considered as low, especially when compared to the Euroarea average (92.9 percent of GDP), (figure 1.7). The average gap between the two regions of around 48 pp is relatively high, indicating another important deficiency in European transition economies. Furthermore, the relatively high level of savings compared to the low level of crediting suggests that deposits are not constraining the banking system to issue more loans. The private sector credit to GDP ratio is lowest in Kosovo, Macedonia and Albania, which may indicate that the banking systems in these countries perceives the real sector as more risky or have fewer profitable projects to lend to. On the other side, even Estonia with the highest private credit to GDP ratio of 63 percent stands significantly below the Euroarea average (see table A1.7 in appendix 1 for numerical data).[5]

---

[5] Private sector loans include loans to households, non-financial corporations and other private sector bodies, but exclude loans granted to government, central banks, financial institutions and public entities.

Figure 1.6. Private sector credit to GDP, in percent



Source: Respective central banks and ECB (2015), own calculations.

1.4. Evidence on the presence of excess liquidity

The analysis of the data in the previous section indicated that banks in European transition economies lend to the private sector at much lower levels than those in advanced economies. Furthermore, their reliance on domestic deposits to finance banks' lending has lessened their external risk, which has contributed to relatively stable liquidity conditions. In order to analyse whether banks in transition economies have sufficient liquid funds to narrow the credit gap, promote growth and hence reduce the output gap, two commercial banks positions will be analysed: their liquidity position at the central bank and their overall liquidity position (liquidity ratio).

The reserve requirement rate (RRR) is a monetary policy instrument aimed at adjusting the banking system's liquidity and stabilizing interest rates. The RRR represents the liquid assets that a central bank mandates commercial banks to keep at all times, usually expressed as a percentage of the bank's total deposits (in local and foreign currency). Commercial banks in many of the European transition economies appear to be keeping reserves in excess of the RRR required by the central bank (EL1 in figures 1.8 to 1.18 below).[6] The required part of

[6] The definition of the concept of 'excess liquidity' is not a straightforward matter in the literature, thus it will be thoroughly elaborated in chapter 4, where a new definition will be proposed. In chapter 1, all the liquid funds kept by commercial above regulatory ratios will be considered as excess liquidity or excess reserves. Later, in

reserves to be held at the central bank is usually remunerative based on the specific regulation of each country, however any above statutory reserves are usually non-remunerated. The remuneration of the required reserves is typically very low (no more than 0.25 percent of the reserves). Besides the liquid funds kept at central banks, available cash above the required rate kept in commercial banks' accounts, placements abroad, investment in securities or domestic government treasury bills are also classified as commercial banks' liquid funds. Different from the above statutory required reserves at the central banks, most of these instruments are remunerated (Ruffer and Stracca, 2006; Khemrraj, 2006). Some of the central banks of European transition economies also impose a mandatory liquidity ratio, which varies from country to country.[7] For example, from 2009 Albania imposed a 20 percent liquid assets to short-term liabilities mandatory ratio on banks; from 2012 Kosovo imposed a 25 percent of liquid assets to short-term liabilities ratio to be held at all times; Macedonia since 2011 imposed ratios of liquid assets to short-term liabilities of the same maturity of 100 percent (similar to LCR, see footnote 7); Serbia has no mandatory liquidity ratio, whereas the rest of the sample countries which are part of the European Union that also implement Basel III framework, since 2013 are uniformly subject to the liquidity coverage ratio (LCR) and Net Stable Funding (NSF) ratio.[8]

Even those countries that have left their banking system free of mandatory liquidity ratio throughout the 2000s have imposed regulatory restrictions on liquidity positions (administrative, structural or exposure restrictions).[9] The part of liquid funds held by commercial banks above the mandatory ratio may be considered as another part of excess

chapter 4, a further disaggregation between precautionary and involuntary excess liquidity will help define whether excess reserves are seen as excess from banks' or from the economy's perspective.

[7] The mandatory liquidity ratio (e.g. liquid assets to short-term liabilities or liquid assets to total assets) was a popular prudential measure in the late '90s and the beginning of 2000s. However, since then many of the regulatory institutions concentrated on capital buffers and credit risk as the most important banking risks. As the liquidity risk was perceived to be one the least important banking risks, mandatory liquidity ratios were gradually abandoned. After the global financial crisis, the new regulation of Basel III reintroduced mandatory liquidity ratios.

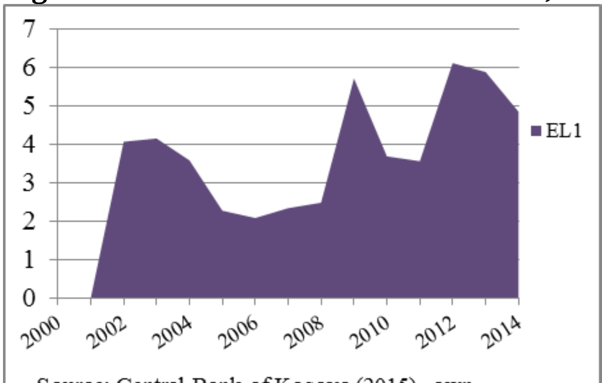
[8] The LCR represents the ratio of stock of high quality Liquid Assets to Net Cash Outflows over 30 days and this should be higher than 100 percent; the NSF is the ratio of Available Amount of Stable Funding to Required Amount of Stable Funding and should also be higher than 100 percent. The LCR aims at promoting short-term resilience, whereas the NSF aims at promoting resilience over a one year time horizon by creating additional incentives for banks to fund their activities with more stable sources of funding on an ongoing basis.

[9] In chapter 4, when defining and measuring excess liquidity in more detail, the mandatory liquidity ratios in less developed European transition economies are set at around 20 percent, whereas in more advanced economies where financial sector is more developed in terms of instruments to invest the mandatory ratio is set to 15 percent. The same ratios are used to calculate the EL2 in chapter 1.

liquidity (EL2 in the figures below). Thus, both parts of liquidity positions, namely EL1 and EL2 are treated as excess in our study.

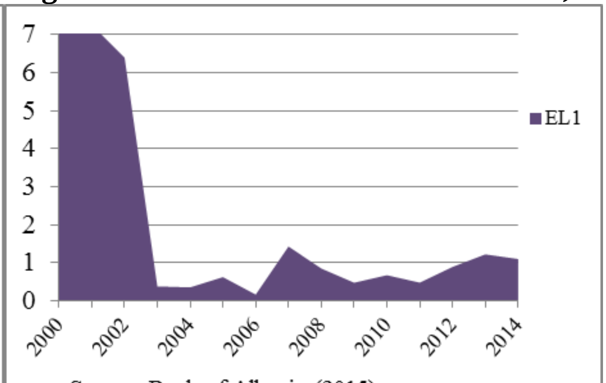
Given the deficiencies of the banking systems in many European transition countries identified in the previous section, we are particularly interested in the additional funds, i.e. excess liquidity, that could contribute to enhancing credit to the private sector and promote growth. Evidence on the presence of excess liquidity held above mandatory ratios in the banking system of European transition economies is presented in the figures 1.7 to 1.17 below.[10]

Figure 1.7. Kosovo: EL1 to Total Assets, %



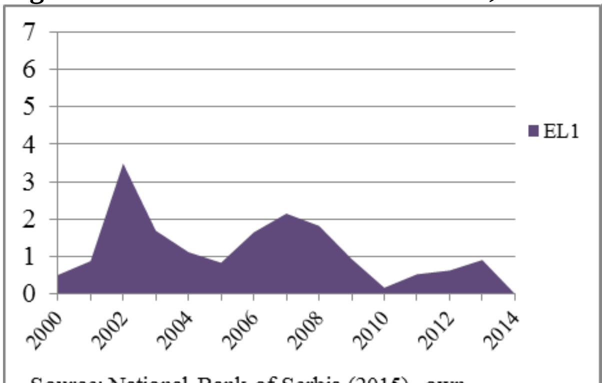
Source: Central Bank of Kosovo (2015), own calculations.

Figure 1.8. Albania: EL1 to Total Assets, %



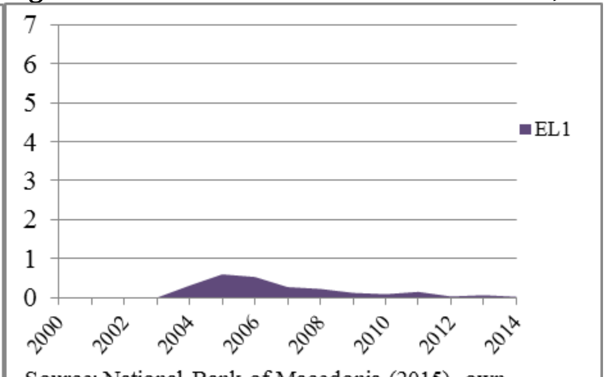
Source: Bank of Albania (2015), own calculations.

Figure 1.9. Serbia: EL1 to Total Assets, %



Source: National Bank of Serbia (2015), own calculations.

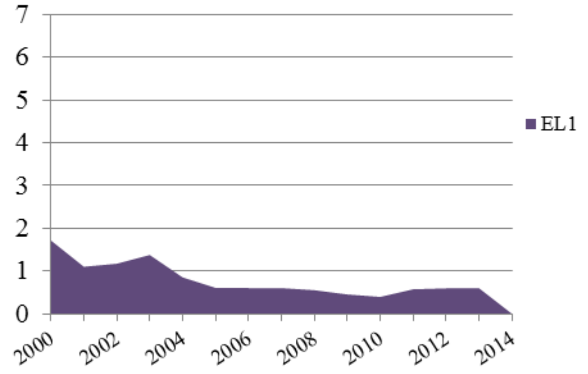
Figure1.10. Macedonia: EL1 to Total Assets,%



Source: National Bank of Macedonia (2015), own calculations.

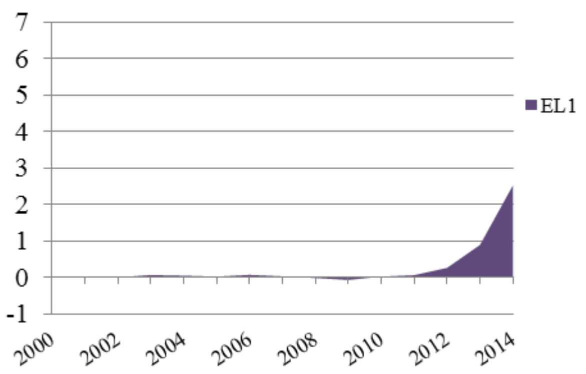
[10] The liquidity data are usually not reported in a consistent form (such are balance sheet data) from different central banks and are not always publicly available. Thus, central bank officials were individually contacted to gain the needed data. Due to the lack of data, some European transition economies are not presented in figures 1.8 – 1.24.

Figure 1.11. Estonia: EL1 to Total Assets, %



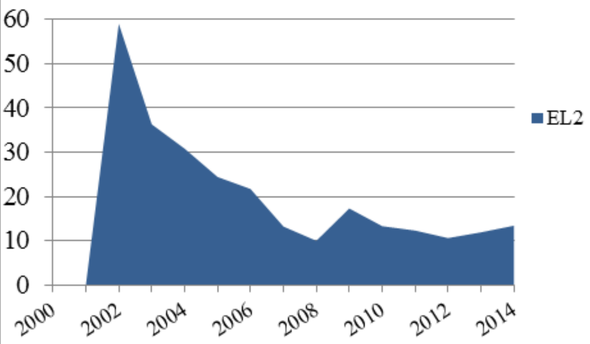
Source: Bank of Estonia (2015), own

Figure 1.12. Latvia: EL1 to Total Assets, %



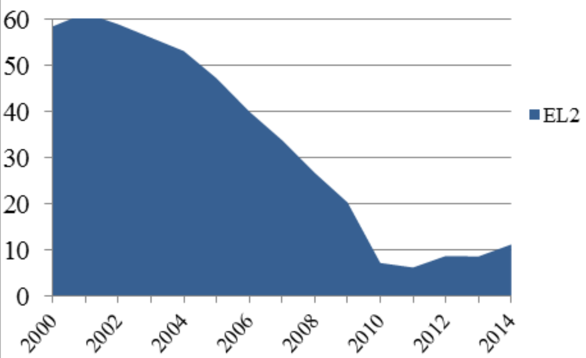
Source: Bank of Latvia (2015), own calculations.

Figure 1.13. Kosovo: EL2 to Total Assets, %



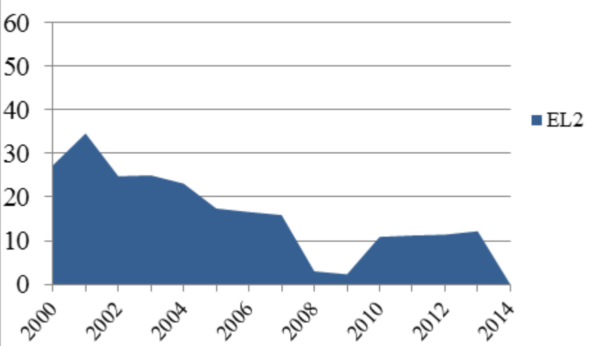
Source: Central Bank of Kosovo (2015), own calculations.

Figure 1.14. Albania: EL2 to Total Assets,%



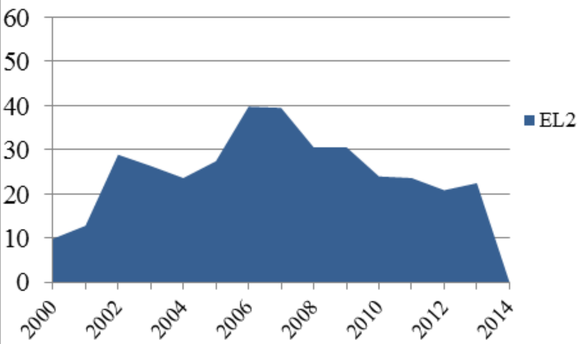
Source: Bank of Albania (2015), own calculations.

Figure 1.15. Macedonia: EL2 to Total Assets, %



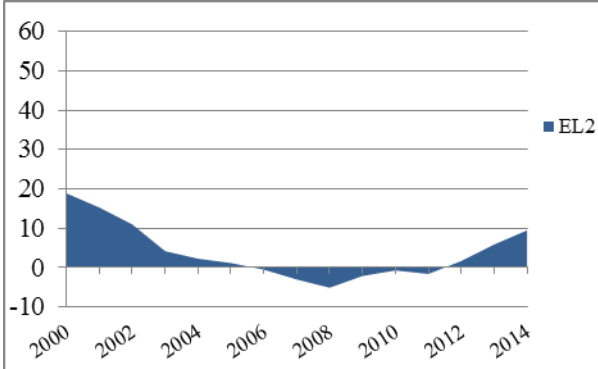
Source: National Bank of Macedonia (2015), own calculations.

Figure 1.16. Serbia: EL2 to Total Assets, %



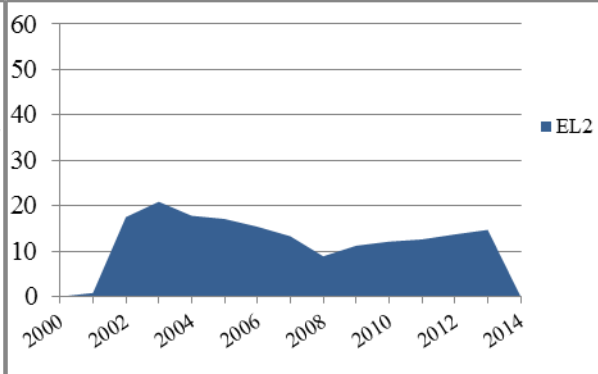
Source: National Bank of Serbia (2015), own calculations.

Figure 1.17. Hungary: EL2 to Total Assets, %



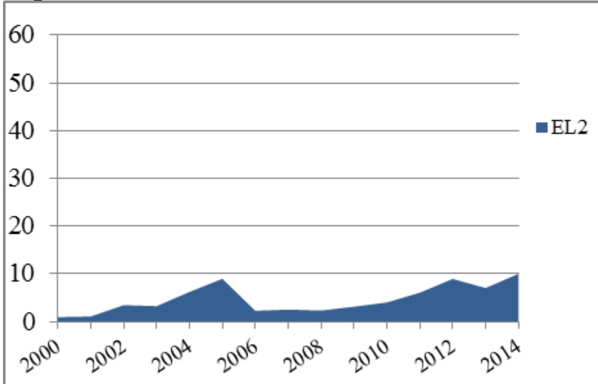
Source: National of Hungary (2015), own calculations.

Figure 1.18. Czech Rep: EL2 to Total Assets, %



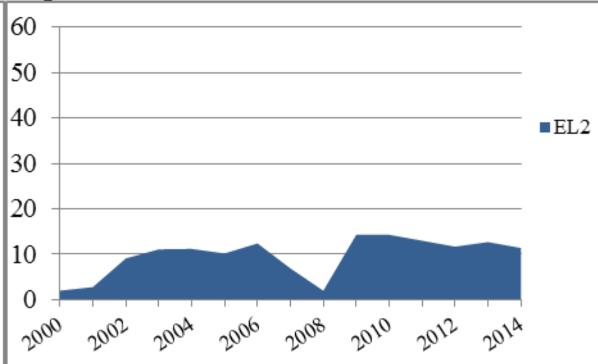
Source: Czech National Bank (2015), own calculations.

Figure 1.19. Estonia: EL2 to Total Assets,%



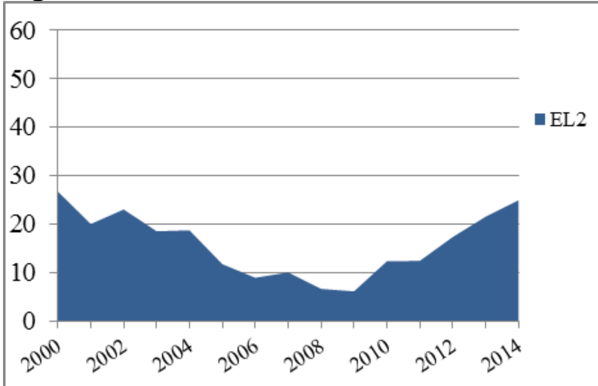
Source: Bank of Estonia (2015), own calculations.

Figure 1.20. Poland: EL2 to Total Assets,%



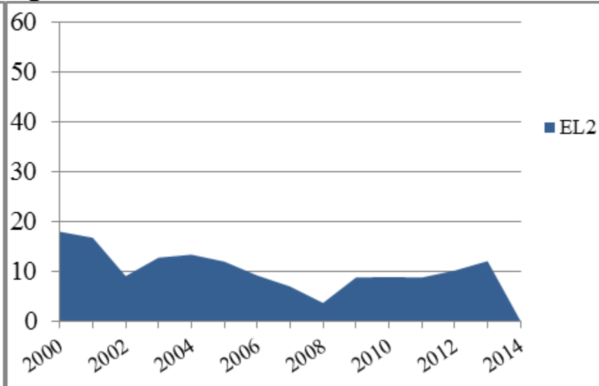
Source: National Bank of Poland (2015), own calculations.

Figure 1.21. Latvia: EL2 to Total Assets, %



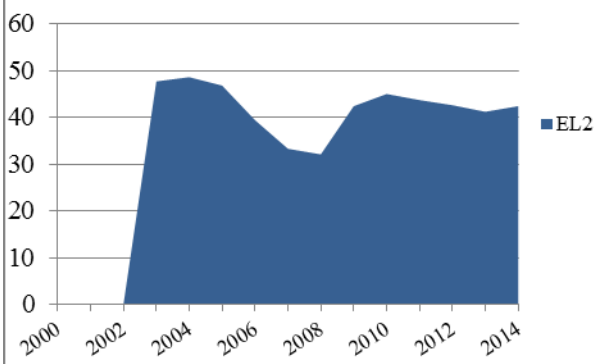
Source: Bank of Latvia (2015), own calculations.

Figure 1.22. Lithuania: EL2 to Total Assets, %



Source: Bank of Lithuania (2015), own calculations.

Figure 1.23. Poland: EL2 to Total Assets, %



Source: National Bank of Poland (2015), own calculations.

An overall analysis of the figures 1.7 to 1.23, suggests that three common characteristics may be observed. First, with the exception of Hungary, all the selected countries have been holding excess reserves above mandatory ratios for the whole observed period (14 years), implying the availability of funds to enhance lending, despite low levels of credit (see section 1.3), even during the time of global financial crisis. Second, a decline of excess reserves is depicted in all the observed countries around the onset of the global financial crisis (2007 to 2009). In the more developed and integrated European transition economies like Hungary, Estonia and Lithuania the decline started earlier (around 2007), whereas in the less developed ones (e.g. Kosovo and Albania), the decline started later (around 2009). Third, there was an increase in excess reserves soon after the crisis in all the selected countries, indicating that the crisis' impact in this respect may have been only transitory. However, the lower level of excess reserves in the aftermath of the crisis may have been a result of lower savings rather than higher loans (see figure 1.5 and 1.6).

In a full-scale euroized economy, like Kosovo, it is the government rather than the central bank that needs to provide a large part of the resources for emergency liquidity assistance to the financial system. However, due to the continuous favourable liquidity position in the banking system the need for liquidity intervention has never materialised. Furthermore, as depicted in figure 1.8, banks in Kosovo have historically kept above than required reserves at central banks (EL1), as well as relatively high liquidity ratios, a part of which is above the mandatory ratio (EL2), (see table 1.8 in Appendix 1 for numerical data). Starting with relatively high levels of excess liquidity at the beginning of 2000s (which corresponds with the establishment of banking system in Kosovo and weak intermediation), soon after and up to 2008, the excess liquidity funds started to decline, as this was a period of robust credit growth. When the global financial crisis started to impact in 2009, it was the restrictions from parent banks regarding lending policies, rather than a higher risk perceived by the banks regarding the economic environment, led to a slower and more conservative lending approach.[11] While lending activity had slowed down, the accumulation of excess liquidity started to rise. Meanwhile, the bank placements abroad also grew more, which was possibly a manifestation of the excess liquidity. The Czech Republic is also one of the selected countries continuously operating in an environment of excess liquidity. Excess liquidity above the reserve requirements at central bank is evident in levels (see table A1.13 in Appendix 1 for numerical data), however as a ratio to total assets is fairly low thus almost invisible in figure 1.19 (i.e. EL1). A systemic liquidity surplus has even made the Czech banking sector (mainly subsidiaries of banks from the Euroarea) capable of serving as net creditor to its parent banks, even during the crisis (IMF, 2012a). The Estonian banking sector has been highly liquid throughout the observed period, and even during the financial crisis operated with excess reserves (figure 1.12, see table A1.12 in Appendix 1 for numerical data). The liquidity cushions prior to the crisis and the support from Nordic parent banks helped in preventing liquidity problems during the crisis, whereas the relatively high reserve requirements further boosted the liquidity position. Therefore, all the countries in the sample had significant excess reserves, which could have been used for issuing new loans, while issuing relatively low levels of credit.

In the two following sections further analysis will continue to investigate whether low levels of credit to GDP and the manifestation of excess liquidity are a consequence of low bank profits, high risk, inefficient decision making or information asymmetry.

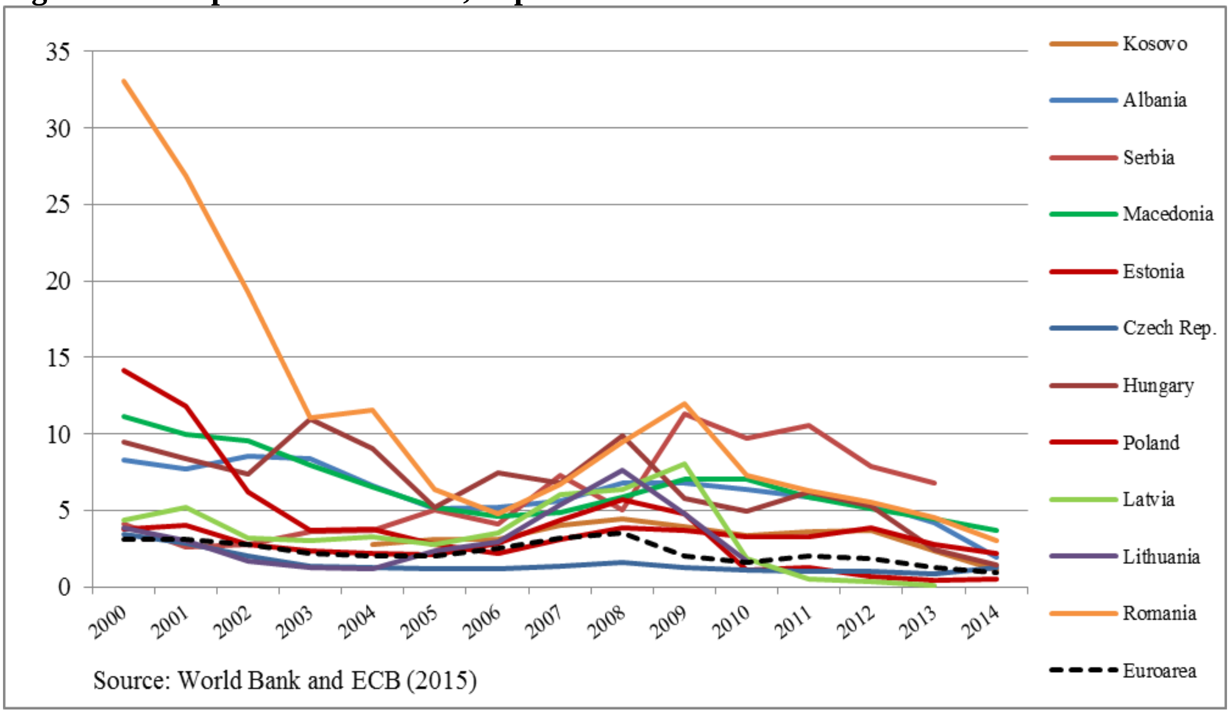
1.5. Banking risks, costs and profitability

After establishing evidence on the presence of excess liquidity in the previous section, this section will elaborate on some of the banking characteristics, market and institutional factors that may impact the lending decision making and hence the level of excess liquidity. The European transition economies have made considerable progress in terms of market liberalization and competitiveness in the banking sector throughout the transition period. The

[11] After the onset of the GFC, due to large losses incurred by the banks in developed European countries, many banks in these countries fell under the capital adequacy ratios (CAR). In order to increase the CAR, most of the banks in these countries (most of them parent banks to subsidiaries in European transition economies) had to reduce the risk-weighted assets (RWA), where credit comprises the largest part. Consequently, bank subsidiaries in transition economies were also under the requirements to reduce the RWA, hence the overall credit to the economy.

deposit interest rates in the observed period stood at an average of 5 percent, with a slightly positive trend from 2007 through 2009 (figure 1.24). “One obvious consequence of the financial crisis in these financial sectors was an increase in deposit interest rates confirming the hypothesis that banks more aggressively looked for deposits from households” (and therefore increasing liquid funds), (Jazbec and Kastrati, 2011, p. 84). In the Euroarea, domestic funding is typically 3pp cheaper than in transition economies (see table 1.19 in Appendix 1 for numerical data).

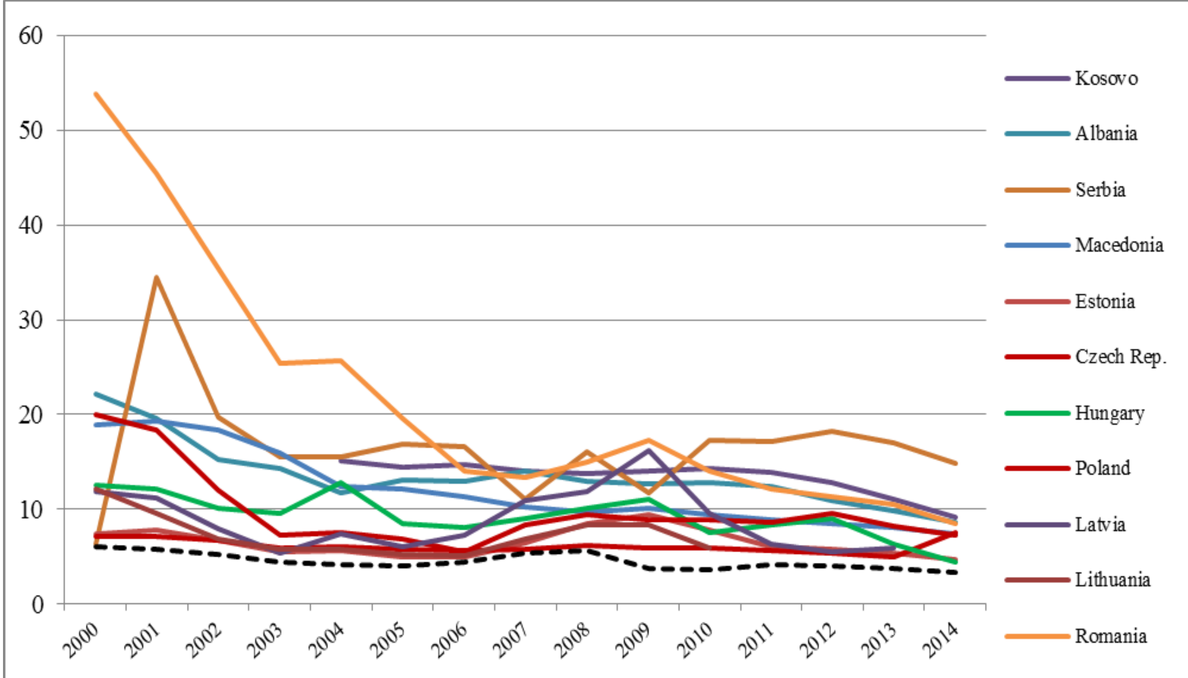
Figure 1.24. Deposit interest rates, in percent



On the other side, high interest rates on loans, which may indicate a lack of competitiveness in the banking system, are considered as an obstacle preventing access to finance in many transition economies (Bernanke et al., 1998; Beck et al., 2005; Inderst and Mueller, 2007). The loan interest rates have remained at relatively high levels, averaging around 11 percent (figure 1.25), compared to an average interest rate of 4.5 percent in the Euroarea.[12] As a result, the average interest rate spread in European transition economies is around 6 pp, compared to the 2.2 pp interest rate spread in the advanced European economies.

[12] The nominal interest rate is equal to the real risk-free rate, plus an inflation premium, plus a default risk premium, plus a liquidity premium, plus a maturity risk premium.

Figure 1.25. Lending interest rates, in percent

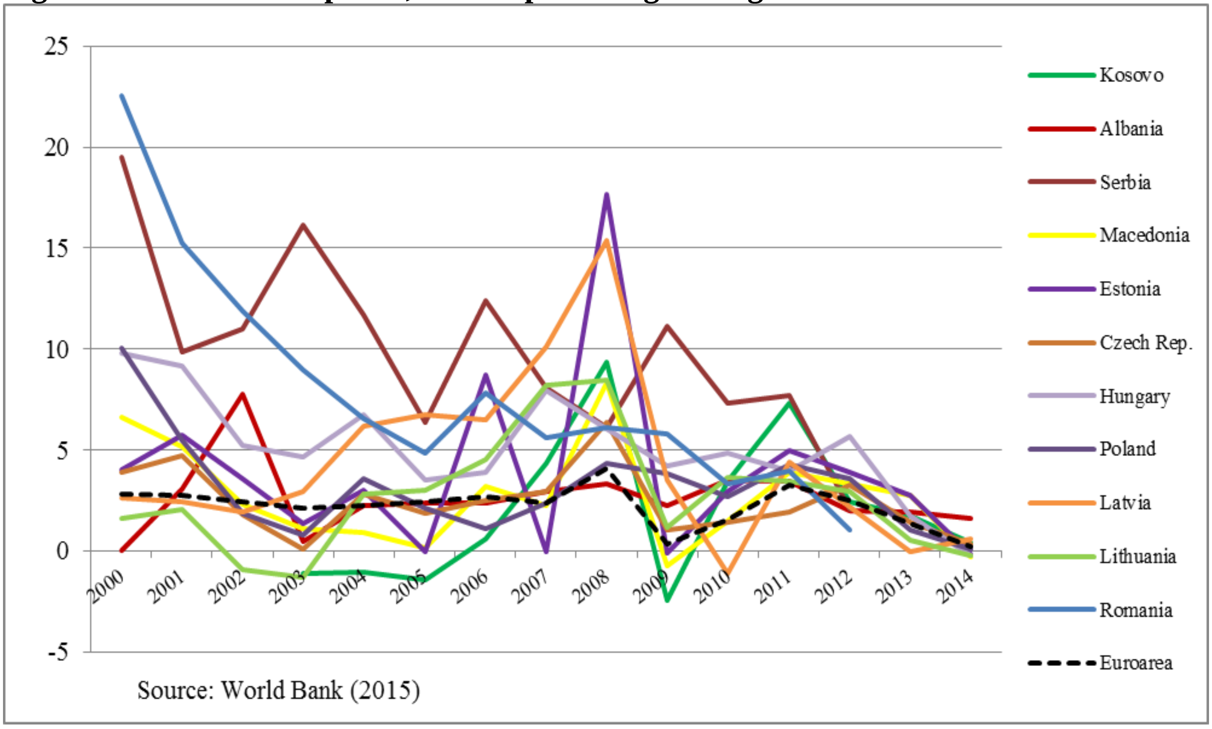


Source: World Bank and ECB (2015)

During the financial crisis, while banks in developed European countries have reduced their lending rates (such as Austria, Germany, Greece etc.), those in European transition economies increased them further due to credit restrictions from parent banks (leading to a lower supply of new credit and a termination of capital flows in some other countries (e.g. Montenegro) and also loan portfolio deterioration (IMF, 2012b). The high interest rate spread in European transition economies mainly reflects higher lending rates Also, although relatively high deposit rates may be convenient for depositors, they represent expensive funds to commercial banks. Therefore, larger interest rate spreads between loans and deposits in transition economies relative to advanced European countries would seemingly only encourage banks to lend more. Despite high lending interest rates, as indicated in section 1.4, banks in these countries still operate with excess reserves. This may imply that above a certain level of loan interest rate some borrowers are being credit rationed, so the accumulation of excess liquidity may be a supply-side phenomenon. On the other side, high lending rates, especially in SEE countries (around 13.9%), may also stem from the higher risks perceived by the financial sector.

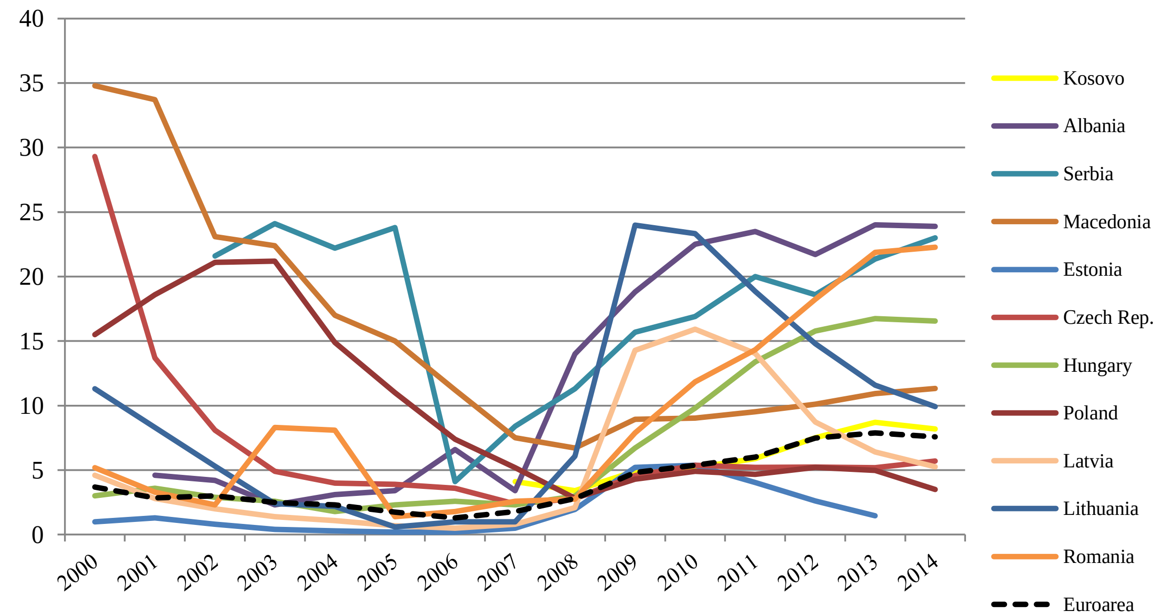
Speaking of banking risks, higher expected inflation represents one of the factors increasing the risk premium in the loan interest rate. The average inflation rate in European transition economies during 2000s (excluding Serbia since it can be considered an outlier) was around 3 percent (figure 1.26; see table A1.21 in Appendix 1 for numerical data). In terms of real interest rates, the average real deposit rates would be around 2 percent, whereas the real lending interest rates would be 8 percent. Therefore, even after adjusting for inflation, the real interest rates in European transition economies may still be considered as relatively high compared to those in the Euroarea. Furthermore, changes in inflation rate prior to and after the GFC did not significantly affect the average lending rates, indicating that it was not a crucial risk factor determining the lending rates.

Figure 1.26. Consumer prices, annual percentage change



One reason for setting high lending rates by the banks may be to compensate for a relatively higher level of non-performing loans (NPLs) as compared to those in the Euroarea countries. The NPL ratio, which is usually considered as a typical risk measure (Berger and DeYoung, 1997; Podpiera and Weill, 2008) in the banking sector, is on average 5pp higher than the ratio in advanced economies of Europe (figure 1.27; see table A1.22 in Appendix 1 for numerical data). However, in a visual depiction of lending interest rates versus NPL rates by country, it can be noted that the correlation between the two is fairly weak, thus the lending interest rates are not necessarily set high to offset the loan repayment risk.

Figure 1.27. Non-performing loans to total loans, in percent



Source: World Bank (2015)

Another reason for setting high lending interest rates may be justified by poor financial performance (e.g. bank losses or low profitability). However, most of the banks in European transition economies are profitable. The highest share of banking system profits is mainly generated by the subsidiaries of foreign owned banks in these economies, which comprise around more than 80 percent of the banking sector. For example, Eastern Europe represents the most important part of the crediting activity of ProCredit Holding (PCBH), a German based multinational financial company. The PCBH's subsidiary is the largest bank operating in Kosovo, alone generating more than 40 percent of the overall group's profits (PCBH, 2014). Raiffeisen Bank International (RBI) is another multinational financial company with banking activities spread throughout the Central and SEE countries.[13] The financial performance of the RBI has been generally positive throughout 2000s, thanks particularly to the positive performance of its subsidiaries in Central and Eastern Europe (CEE) and SEE Europe, which generate more than 60 percent of the RBI's profit (table 1.2). Therefore, despite high funding costs and NPL rates, the high interest spreads do not seem to be associated with high risks, but rather reflect the high profitability of the banks in these economies. If this is the case, the accumulation of excess liquidity (see section 1.4) in banks

[13] The RBI operates in CEE (Czech Republic, Hungary, Slovakia, Slovenia and Poland) and SEE countries (Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Kosovo, Romania and Serbia).

of European transition economies appears to be even more unusual and interesting to investigate.

Table 1.2. The share of SEE and CEE countries' profit in total RBI Group profit, in percent

Country/ Year	2005	2008	2011	2013	*Average*
SEE Countries	40.7	41.7	36.0	45.9	*41.1*
CEE Countries	50.0	32.6	3.4	10.8	*24.2*
Other	9.3	25.7	60.7	43.3	*34.7*
RBI Group profit (in mil. Euros)	460.0	1,078.0	976.0	603.0	*779.3*

Source: RBI Annual Reports (2005, 2008, 2011, 2013)

The following section presents further evidence on institutional and legal conditions that may facilitate lending procedures in European transition economies, as well some additional evidence from the firms' perspective on loan applications.

1.6. Constraints on bank lending in European transition economies

Despite the banks' good health and ample liquidity, the loans to GDP ratio in European transition economies remains relatively low, partly because of the environment in which the banks operate. The slow economic revival and lower than expected growth rates in the region cannot be attributed to the macroeconomic developments and fragile financial systems in those countries alone (World Bank, 2011). Yet, a crucial point is that these economies also lack the institutional setting which would support the investment demand needed to generate the economic growth necessary for a speedy catch-up with the more advanced European countries. The assumption of efficient financial intermediaries reducing asymmetric information and lowering the transaction costs of investments (Levine, 1997) may not always hold for transition countries. This is because, transition countries in general, and SEE ones in particular are considered to have a set of problems, such as inefficiencies in banks themselves, poorly functioning institutions, lack of the rule of law implementation and other institutional and market failures. Due to these economy's high levels of informality and their inefficient legal systems, interest rate spreads are high, and banks apply high collateral requirements. Thus, another set of banking risks that may be associated with lower levels of credit in European transition economies than in Euroarea, are the higher risks associated with the weak legal protection of banks and other financial institutions in enforcement of existing business contracts.

Table 1.3. Ease of doing business indicators, rankings 1 - 189 (1 = best, 189 = weakest)[14]

Country/ Year	Getting credit 2006	Getting credit 2009	Getting credit 2011	Getting credit 2014	Paying taxes 2006	Paying taxes 2009	Paying taxes 2011	Paying taxes 2014	Enforcing Contracts 2006	Enforcing Contracts 2009	Enforcing Contracts 2011	Enforcing Contracts 2014
Kosovo	…	…	15	23	…	…	149	63	…	…	89	138
Albania	48	12	32	36	125	143	41	131	99	89	155	102
Serbia	33	28	46	52	64	126	33	165	76	96	65	96
Macedonia	48	43	32	36	79	27	30	7	72	70	50	87
Estonia	48	43	32	23	79	34	30	24	72	30	50	16
Czech Rep.	21	43	46	23	75	118	128	119	1	95	78	37
Hungary	21	28	32	17	118	111	109	88	8	12	22	20
Poland	65	28	15	17	71	142	121	87	112	68	77	52
Latvia	13	12	6	23	52	36	59	24	11	4	14	16
Romania	…	12	15	7	…	146	151	40	…	31	54	51

The getting credit indicator assesses the legal rights of borrowers and lenders in secured transactions.

Paying taxes indicates the taxes and mandatory contributions that a medium-size company must pay in a given year as well as measures of the administrative burden of paying taxes and contributions

Enforcing contracts measure the efficiency of the judicial system in resolving a commercial dispute.

Source: World Bank (2015)

The World Bank Doing Business Reports (2006 - 2014), indicate that environmental factors, outside of banks' control, in the transition countries in most of cases have a low ('weak') ranking, relative to other countries. For example, the enforcing contract indicator ranks relatively low in most of the European transition economies, with the lowest ranks during recent years in SEE countries (table 1.3). Weak contract enforcement may induce moral hazard behaviour amongst borrowers and lower the bank's expected returns, therefore the bank may lend less. The paying taxes indicator may be considered as one of the crucial factors in determining the banks' decision in issuing a loan, as tax declaration documents validate the turnover, income capacity and the debt situation of a firm. The relatively low ranking of the selected countries regarding tax payment may suggest that the poor tax payment by companies may impact the banks' decision whether to approve loan applications. Due to these and other problems in tax reporting and contract enforcement, getting credit may also be an obstacle for firms in these countries. From three selected indicators in table 1.3, getting credit appears to be less of an obstacle, implying that firms have an easier access to credit funds. However, for the calculation of 'getting credit', indicators are weighted proportionally, based on their contribution to the total score, i.e. 60 percent assigned to the strength of legal rights index and 40 percent assigned to the depth of credit information index. Therefore, the high rankings presented in table 1.3 may not reflect the realities facing would-

[14] The ranking of each country in respective indicators is relative to other countries. "The ranking calculation is set between two frontiers, best (based on best practices) and worst (for worst practices)" (Doing Business, 2015, p. 25).

be borrowers in these economies. This is because, although these economies do well in terms of laws and availability of information on borrowers (credit registries), this may not mean much in terms of the availability of credit to businesses.

Overall, these market imperfections that characterize the environment within which the banking sector of European transition economies operate may give rise to larger asymmetric information problems between the banks and the borrowers, compared to those in the more developed European countries (which are assumed to have a better functioning institutions). Therefore, besides the banks' optimizing behaviour, as discussed in the previous section, relatively high interest rates may also prevail to offset the costs relating to problems associated with greater asymmetric information and weaker institutions. Another World Bank indicator, "the strength of legal rights index, measures the degree to which collateral and bankruptcy laws protect the rights of borrowers and lenders and thus facilitate lending" (Doing Business, 2015, p. 126). Higher scores indicate greater protection. The average index (table 1.4) of 7.3 out of 12 indicates (Euroarea average 5.8) suggests that the laws designed to facilitate the lending/borrowing process in European transition are broadly adequate, though the implementation and enforcement of these laws may lag behind (EBRD, 2005).

Table 1.4. Strength of legal rights index (0 = weakest, 12 = strongest)

Country/ Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	*Average*
Kosovo	…	…	…	…	…	7	7	7	7	8	8	*7.3*
Albania	9	9	9	9	9	9	9	9	9	8	7	*8.7*
Serbia	6	6	7	7	7	7	7	7	7	5	5	*6.5*
Macedonia	6	6	6	6	6	6	6	6	6	6	6	*6.0*
Estonia	6	6	6	6	6	6	7	7	7	7	7	*6.5*
Czech Rep.	7	7	7	7	6	6	6	6	6	5	7	*6.4*
Hungary	7	7	7	7	7	7	7	7	7	6	10	*7.2*
Poland	8	8	8	8	8	9	9	9	9	7	7	*8.2*
Latvia	10	10	10	10	10	10	10	10	10	9	9	*9.8*
Lithuania	5	5	5	5	5	5	5	5	5	6	6	*5.2*
Romania	8	8	8	9	9	9	9	9	9	10	10	*8.9*
Euroarea	6	6	6	6	6	6	6	6	6	5	5	*5.8*

Strength of legal rights index measures the degree to which collateral and bankruptcy laws protect the rights of borrowers and lenders and thus facilitate lending.

Source: World Bank (2015)

Table 1.5 below presents the depth of credit information index, which helps banks in revealing personal information regarding (potential) borrowers and their credit profile (rankings) to evaluate the loan repayment capability, which information directly impacts the decision making on whether or not to issue a loan. On average, this index ranks European transition economies higher than those in the Euroarea, indicating the availability of credit information to banks to facilitate lending is relatively good. This finding does not appear to be consistent with the credit rationing argument (Waller and Lewarne, 1994) that high lending rates are primarily due to high asymmetric information problems in these economies.

Table 1.5. Depth of credit information index, 0 = low to 8 = high

Country/ Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average*
Kosovo	…	…	…	…	…	5	5	5	5	6	6	5.3
Albania	0	0	0	0	5	5	5	5	5	6	6	3.4
Serbia	0	0	4	4	4	5	5	5	5	7	7	4.2
Macedonia	0	0	0	0	4	4	4	6	6	7	7	3.5
Estonia	5	5	5	5	5	5	5	5	5	7	7	5.4
Czech Rep	4	5	5	5	5	5	5	5	5	7	7	5.3
Hungary	0	0	5	5	5	5	5	4	4	5	5	3.9
Poland	4	4	5	5	5	6	6	6	6	8	8	5.7
Latvia	0	0	0	0	0	4	4	4	4	5	5	2.4
Lithuania	0	4	4	6	6	6	6	6	6	8	8	5.5
Romania	0	0	4	4	4	5	5	5	5	7	7	4.2
Euroarea	3.6	4.0	3.7	3.9	3.7	4.0	4.1	4.0	4.1	5.6	5.7	4.2

*Note: Depth of credit information index measures rules affecting the scope, accessibility, and quality of credit information available through public or private credit registries.

Source: World Bank (2015)

Finally, table 1.6 below presents a demand side perspective on loan application from firms. From all the responding firms in BEEPS questionnaire in selected European transition economies in years 2002, 2005 and 2009, only around third had applied for a loan. The relatively high proportion of firms not applying for a loan may serve as an explanation for the low level of credit to GDP in this region (see section 1.3), as banks may unwillingly hold other assets when the demand for loans may be low. Otherwise, having in mind the relatively high lending interest rates as compared to those in Euroarea (see section 1.5) and Hashi and Toci’s (2010) findings, many of those firms who have not applied may also represent discouraged borrowers due to non-price mechanisms, such as high collateral requirements.

Table 1.6. Firms that applied for a loan in the last fiscal year, in percent

Country/	2002		2005		2009	
Year	yes	no	yes	no	yes	no
Albania	23.8	76.3	28.9	71.1	34.7	65.3
Poland	28.1	71.9	39.4	60.6	21.4	78.6
Romania	36.0	64.0	37.9	62.1	24.1	75.9
Serbia	33.8	66.2	25.8	74.2	52.5	47.5
Moldova	30.6	69.4	40.0	60.0	37.0	63.0
B&H	26.2	73.8	56.0	44.0	29.7	70.3
MKD	20.5	79.5	…	…	22.9	77.1
Estonia	29.7	70.3	…	…	24.6	75.4
Czech	31.4	68.6	…	…	21.1	78.9
Hungary	23.8	76.2	…	…	22.2	77.8
Latvia	25.4	74.6	…	…	16.3	83.8
Lithuania	33.5	66.5	…	…	44.1	55.9
Slovakia	43.1	56.9	…	…	48.8	51.2
Slovenia	49.4	50.6	…	…	49.4	50.6
Bulgaria	31.9	68.1	…	…	46.2	53.8
Croatia	20.9	79.1	…	…	41.9	58.1
Monenegro	…	…	…	…	50.0	50.0
Average	*30.5*	*69.5*	*38.0*	*62.0*	*33.5*	*66.5*

*Note: Question k16: 'Referring again to the last fiscal year, did this establishment apply for any loans or lines of credit?' The calculated average of all the responding firms in respective years excludes firms with no answer (in BEEPS questionnaire coded as '-9'). Due to missing responses for most of the sample countries, year 2007 was not presented.

Source: Own calculations based on BEEPS Panel Questionnaire (2009)

To conclude, it appears that even after more than a decade of the transition process these economies still need to strengthen their business environment by taking actions to improve contract enforcement and create a broader tax payment base to help in extending credit growth and broadening their firms' access to credit (Roaf et al., 2014).

1.7. Three sample countries – the Czech Republic, Estonia and Kosovo

The macroeconomic imbalances and structural and market deficiencies identified in section 1.3 led us to believe that many European transition economies are operating significantly below potential capacity. In order to be able to assess the below potential activity of these economies, some quantifiable indicators will need to be estimated in order to conduct a deeper analysis. These analyses usually consist in generating the business cycle of a country, which implies the decomposition of the actual output into trend (potential output) and cycle (output gap). Given that these two indicators are unobservable by nature, they need to be independently estimated by researchers. Furthermore, business cycle analysis usually represents a country by country exercise, thus the potential output and the output gap indicators need to be separately estimated for each sample country. As will be seen later in chapters 2 and 3, among several estimation techniques, we will choose a relatively complex and advanced method: the unobserved method operationalised via the Kalman filter. This approach in terms of transition economies with relatively short time-spans and structural changes (Kalotay, 2001; Filer et al., 2001; Orts et al., 2008), can be very time consuming and challenging to generate 'reasonable'[15] results. Therefore, instead of implementing a repetitive exercise for all European transition economies, which may not necessarily add value to the analysis, three sample countries to represent the European transition region were chosen and a more thorough analysis for each country was conducted. The chosen countries are the Czech Republic, Estonia and Kosovo.

The reasons for choosing these three specific countries are twofold. Firstly, given that within the European transition economies the level of economic development differs by region, the three selected countries are chosen to represent different levels of development and region. The Czech Republic represents the northern part of European transition countries, which is also economically more developed (including for example nearby countries like Poland and Hungary); Estonia is less developed economy than the Czech Republic, representing the Baltic countries as well as one of the countries severely hit by the Global Financial Crisis; while Kosovo represents one of the least developed countries, also representing the SEE countries. Secondly, the selected countries are also representative of the diversity of economic structure and evolution post-transition.

1.8.The key research questions

Based upon the context outlined above in sections 1.2 to 1.6, this research project initially proposed several research questions for investigation. After identifying some of the macroeconomic deficits and gaps in section 1.3 that most of the selected transition countries share, the first research question raised was: are European transition economies operating with large output gaps and if yes how large are their output gaps? Even though there are some

[15] By 'reasonable' results the author implies that the estimated model should initially converge (with transition economies' data very often it does not even after a considerable times of iterations), the series should be free of serial correlation after a certain lag and the generated (estimated) series should approximately represent the conventional form of the business cycle.

estimates of potential output and output gaps for the European transition economies (Room, 2001; Zaman, 2002; Darvas and Vadas, 2003; Benes and Ndiaye, 2004; Antonicova and Hucek, 2005; Tsalinski, 2007), these are mainly focused on their relevance with regard to monetary and fiscal policy conduct. The originality of our research question stems from the concern in identifying common causes of high output gaps throughout the European transition economies. Based on the arguments in section 1.3 and 1.4 indicating low levels of credit to GDP, with excess liquidity in their banking systems for more than ten consecutive years, while their lending interest rates are more than double those in Euroearea, the second research question of the thesis was: why do banks in European transition economies continuously operate with excess liquidity in the face of seemingly profitable opportunities for increasing loans? Further, while addressing the second research question, a set of other sub-research questions arose: what is excess liquidity, what are its main determinants, and is excess liquidity a predominantly a demand or supply side phenomenon in transition economies? The redefinition of excess liquidity concept in this thesis is novel, as well as the overall investigation of this phenomenon which has not previously been conducted for European transition economies. After investigating the first two sets of research questions, the last part of the thesis will be concerned with the question: is excess liquidity held by banks dampening growth and causing larger output gaps? In other words, is the relationship between the output gap and excess liquidity in European transition countries, causal, simultaneous or otherwise endogenous. The endeavour to provide a theoretical and empirical relationship between output gap and excess liquidity is entirely novel, as to the best of our knowledge this particular relationship has not been previously investigated for these economies. Another research question initially planned to be addressed in this research project had to do with the micro-level output gap (Graff and Sturm, 2010; Alichi, 2015), i.e. the determinants of capacity utilisation of the firms. However, as the research needed to answer the earlier research questions was greater than originally planned, this last question was eventually dropped. After theoretical and empirical analyses, some of the other original research questions evolved and changed. The evolution of the initially posed research questions will be discussed later in the thesis.

1.9. How the research questions are addressed

Previous analyses of the output gap and excess liquidity will be critically examined in order to develop appropriate definitions for transition economies. Most of the previous studies have explored the estimation of the output gap taking into account theories that apply to developed countries. Therefore, a considerable part of the empirical analysis in this study will critically assess previous techniques for estimating output gap in developed and developing countries, from where the most appropriate method to represent the characteristics of our sample countries will be chosen. Given that the output gap is an unobservable variable, the output gap for the three sample countries, namely the Czech Republic, Estonia and Kosovo will be estimated with the unobserved components method operationalised via the Kalman filter. This model will be estimated using a time series approach with macroeconomic quarterly data. Comparisons will then be made between this estimate and two other common methods (time-series univariate and multivariate approaches). The estimation of potential output and the output gap represents the first analysis of the business cycle in Kosovo, as well as the first use of the Kalman filter method. In order to address the objectives of this research programme two other econometric analyses will be undertaken. A second model will be built using bank-level data from Bankscope. The use of the Bankscope data for excess liquidity analysis is novel in the literature, as well as the sample countries used for this analysis. This model will focus on movements of excess liquidity across different transition economies and over time. The second stage estimations of this model will identify if excess liquidity is a predominantly supply-side problem due to the precautionary motive or a demand-side phenomenon due to the absence of sufficient profitable projects in transition economies. The bank-level model will include most of the European transition economies and will be estimated using the fixed effects approach with AR(1) residuals in a panel framework. A third model will investigate the macro-level relationship between the output gap and excess liquidity. This model will be estimated using quarterly data with seemingly unrelated regressions method. The main sources of data will be the respective national institutions such as central banks and statistical offices, Bankscope, IMF, World Bank and ECB.

1.10. Conclusion

The aim of this chapter was to establish evidence of output gaps and financial underutilised resources; based on which the initial research questions to be addressed in this research programme were formulated. The chapter has presented evidence on output gaps via macroeconomic imbalances and calculations of banking excess liquidity for most of the European transition economies. Further, other peripheral factors, some outside of banks' control, were presented as additional arguments that may impact banks' decisions on lending and hence holding excess liquidity. All the evidence built in this chapter helped in formulating research hypotheses to be tested later in the thesis. Most of the thesis will be concentrated on the three selected European transition countries (the Czech Republic, Estonia and Kosovo), hence a brief economic summary for each country was presented.

The originality of the thesis is threefold: firstly, we provide new estimates of business cycle in Czech Republic, Estonia and Kosovo, along the way explaining the technicalities underpinning the Kalman filtering procedure. Secondly, another theoretical contribution to knowledge is made with regard to the conceptual definition of excess liquidity. Thirdly, a further theoretical and empirical contribution to knowledge concerns the conceptual definition of the relationship between output gap and excess liquidity, which in effect dictates the use of an econometric method, not previously used in similar analysis, in order to investigate their particular relationship.

The structure of the rest of thesis is organized as the following. The second chapter of the thesis elaborates the theory of business cycles and the importance of generating estimations of the unobserved components, namely potential output and the output gap. After a theoretical assessment, potential output and the output gap will be estimated in the third chapter for the Czech Republic, Estonia and Kosovo. The fourth chapter investigates a separate but related issue, that of the definition and determinants of excess liquidity in European transition economies and further distinguishes between precautionary and involuntary excess liquidity. The fifth chapter of the thesis is devoted to an exploration of possible theoretical and empirical links, causal, simultaneous or otherwise endogenous, between the output gap and excess liquidity. The sixth and last chapter of the thesis summarizes the main findings and contributions to knowledge, the limitations of the research undertaken and the potential for future research. Utilising the findings from the previous analyses, the last chapter also considers policy recommendations for transition economies seeking to address high output gaps, specifically targeting the potential gains from the activation of excess liquidity.

Chapter 2. The Output Gap in European Transition Economies

2.1. Introduction

As discussed in the previous chapter, one of the main research questions of the thesis relates to identifying output gaps in European transition economies. In order to address this question initially two business cycle concepts, namely potential output and the output gap are discussed in the European transition context. The theory underlying the determinants of potential output and the corresponding output gap concepts has a long history in economics. In effect, the concept of potential output and the corresponding output gap will initially be presented in the light of mainstream theories, i.e. New-Keynesian and New-Classical theories. However, theoretical and empirical models investigating business cycles are primarily constructed to address developed countries' characteristics and after the Global Financial Crisis even the mainstream theories are being challenged by the need to develop insights into recent economic experience.[16] Hence a specific section discusses the relevance and applicability of potential output and the output gap in transition economies. Given that potential output cannot be directly observed, the output gap is also unobservable. Therefore, before moving to the empirical analysis, a discussion in this chapter will be devoted to choosing an appropriate mechanism for estimating these two indicators in transition economies. While doing so, initially a historical evolution of the filtering methods will be presented, based on the conceptual evolution of the economic trend (potential output). Further, in order to better describe the behaviour of the observed actual output or other macroeconomic time series, commonly used decomposing techniques were taken into account, such as the production function, the SVAR and the Hodrick-Prescott filter. The empirical analysis will largely focus on the description of the unobserved components (Kalman filtering technique) technique, as our chosen method to estimate potential output and the output gap.

The rest of the chapter is organized as follows. Section 2.2 outlines the theoretical framework upon which the business cycle rests. Section 2.3 concerns with the relevance of potential output and the output gap in the transition context. Section 2.4 elaborates several estimation techniques, focusing on univariate and multivariate methods. Section 2.5 concludes the chapter.

[16] The criticisms addressed to the mainstream New-Keynesian model after the onset of the Global Financial Crisis will be discussed in chapter 5.

2. 2. The definition of the potential output and output gap

The *business cycle* in a broader sense represents a recurring economic activity that passes through phases of expansion and contraction. *Potential output* can be defined as the output level generated when the factors of production (labour and capital) are fully employed at the current state of technology. Full employment of the production factors is understood as the maximum production that does not generate inflationary pressures (Okun, 1962), i.e. the sustainable rate of employment. In the short term, if actual output is below or above potential output, it provides indications regarding the inflationary pressures that come from imbalances between demand and supply. In the long-run, potential output indicates the sustainable economic growth position of economy that generates non-inflationary pressures (Room, 2001). The first part of the definition, which refers to the full employment of inputs, has its origins in the growth-accounting theories (Solow, 1957), whereas, the second part of the definition which implies that full employment of factors of production should be understood as some non-inflationary level of input emanates from the inflation-unemployment trade-off theories developed from the Phillips' curve. The essence of growth theories lies in the specification of a production function, which represents the relationship between inputs of production and the aggregate output. In its simplest form, the production function shows how the factors of production, labour and capital and their technological status, generate the level of the output in the economy. For example, in the Cobb-Douglas form of the production function with its exponents summing to unity, the increase in the inputs leads to a proportionate increase in the output; i.e. there are constant returns to scale. The technological state stands for the quality or the productivity of the factors of production. Since it is hard to determine how much of the technological progress refers to labour and how much to capital, the Cobb-Douglas specification defines an overall indicator popularly named as total factor productivity (TFP). In the Cobb-Douglas specification TFP is an unobservable component which has to be estimated via the so called Solow residual.[17] From here, the potential output will be the level of output produced when the factors of production are fully employed and when the total factor productivity is at its 'normal' trend level.[18]

[17] The Solow residual is the difference between the output and the sum of the labour and the capital inputs weighted by the respective shares of labour and capital in output.

[18] Capital input is at its full employment level when the entire available capital stock in the economy is used. The full employment level of the labour input is usually defined as the level of unemployment consistent with low and stable, non-accelerating inflation. The latter is known as non-accelerating inflation rate of unemployment (NAIRU) and it originates from the inflation-unemployment trade-off theory.

The NAIRU concept originates from the inflation-unemployment trade-off theory, firstly developed by Phillips (1958), namely the Phillips curve, which explains the trade-off between inflation and unemployment rate. In this framework, full employment of the factors of production is referred as one consistent with stable inflation and was initially termed the natural rate of unemployment.[19] In this case, the economy is generating its potential output or natural level of output. In the short run, through (unanticipated) changes in aggregate demand the government could stimulate output, employment and prices. In long-run, though, ever accelerating inflation or unemployment permanently above NAIRU are economically and politically unsustainable.

Okun’s law relates changes in unemployment to changes in output or the output gap (Cevik et al., 2013). Given that inflation is believed to be effective in reducing unemployment in the short-run, combining the Phillips curve and Okun’s law would represent a short-run aggregate supply curve. The information on this empirical relationship would also be useful for policymakers regarding the sustainable growth rate. “The ’gap version states that for every one percent increase in the unemployment rate, a country's GDP will be roughly an additional two percent lower than its potential GDP.[20] The ‘difference version’ describes the relationship between quarterly changes in unemployment and quarterly changes in real GDP” (Akram et al., 2014).[21]

The output gap represents the difference between the actual output and potential output. The output gap, defined as the percentage deviation of an economy’s total product from its potential level, indicates the position of the economy in the business cycle. There are two types of output gap: positive output gap and negative output gap. A positive output gap which occurs when the actual output is more than full-capacity output and is usually manifested with excess demand pressures. A positive output gap indicates that the actual output is higher than what can be supported by existing labour and capital resources, i.e. the economic resources will be overworked, and is a leading indicator of changes in the rate of inflation. It will also tend to be associated with a worsening of the current account as consumers buy more imports due to domestic supply constraints. Due to the higher demand for imports, a positive output

[19] The Natural Rate of Unemployment is a hypothetical concept that assumes that in the long run, even in the state of full employment of factors of production, there will always be a non-zero rate of unemployment due to mismatches in the labour market.
[20] Models where the relationship between unemployment and output is specified in deviations from a trend.
[21] Models where the relationship between unemployment and output is specified in terms of differences.

gap may also redirect production planned from exports to the home market. On the other hand, a negative output gap occurs when actual output is less than full-capacity output, i.e. the economic resources will be underworked. This is also called a deflationary (or recessionary) gap, because it will typically cause lower inflation or even deflation.

Following this initial examination of the definition of these two concepts, elaborated mainly in the light of the mainstream theories, the following section will discuss their relevance in terms of policymaking and whether they represent useful and applicable indicators in the context of European transition economies.

2. 3. Is the output gap a relevant concept in transition economies?

Before discussing the theory behind the determinants of potential output and output gap, the relevance of these concepts for transition economies should firstly be established. Although these two concepts are central to much policy making (e.g. inflation targeting and fiscal expenditure planning), they have been mainly discussed in the light of mainstream theories (e.g. Okun's law, the Phillips curve and the production function), these theories are mainly constructed based on the economic conditions in developed countries. Conventional theory defines the business cycle as fitting a mean trend to the actual output, while the business cycle, as proxied by the output gap, will over time symmetrically fluctuate above or below that mean. The standard definition of the business cycle is that an output gap is present when economic activity is either above or below the potential output, caused typically by variations in actual output. In the context of transition economies, the concept of output gap is also subject to large and ongoing structural changes and adjustments, reflected in high structural unemployment rates, imported inflation and widespread market failures that continuously affect and typically limit the potential output. In transition economies, therefore, the output gap may reflect not only the changes in actual output but also in potential output. This issue, coupled with the typically poor quality and short time-span of data make standard theories and empirical approaches non-applicable.

Some of the reasons for the lack of previous studies of the output gap in transition economies may lie primarily in the past economic development of these countries. For example, many transition countries have undergone and are still undergoing a reform process, with the aim of building a market-oriented economy with functional institutions (Svejnar, 2002, Estrin et al., 2009; Roaf et al., 2014). Amongst the first problem to be overcome was the sudden and mass obsolescence of the existing capital stock, especially in the early years of transition due to the old technology prevailing in existing enterprises, in the context of changes in relative prices, whilst the neglect of capital depreciation in the central planning system led to a large scale write down of the value of the existing capital stock (Pistor et al., 2000; Berger et al, 2001; Pyo, 2008 and Roafet al., 2014).[22]

In the context of ongoing reforms, the growth path and the growth variations in transition countries are usually characterized by frequent and large structural breaks and shifts (Jones and Olken, 2005), making research into the business cycle and output gap in these countries during this period extremely complicated or even irrelevant (Jagric and Ovin, 2004). This may particularly be the case in the less developed European transition economies, such as the South Eastern European (SEE) countries. These countries are some of the late comers to the transition process, experiencing a switch in regimes usually in the early '90s and then, during 1997 - 2001, undergoing a period of large reversals in growth patterns (shifts in the growth path or actual output level). For example, in 1997 Albania experienced widespread riots which shifted the growth path and Macedonia's internal conflict in 2001 has also been argued to have permanently shifted the actual and potential output trends (Petkovska and Kabashi, 2011). Even in the more developed European transition countries, like the Czech Republic and Estonia, which nowadays have almost caught-up with the western European Union (EU) countries, some reversals in growth rates have been identified during the 1990s. Therefore, actual output, potential output and the output gap in transition countries may be expected to experience large fluctuations and shifts, rather than conform to the regular business cycle pattern (Neumeyer and Perri, 2004). It may therefore be inferred that transition countries are unique in their economic experiences and thus it will be more challenging to conduct business cycle analysis research in these countries.

The output gap in transition economies has to do with structural as well as the cyclical effects. That is, while measuring the cycle of an economy, such as the aggregate movement of actual output towards expansion or contraction, in transition economies the 'cycle' also reflects structural effects on potential output that may be reflected in the persistent underutilisation of

[22] Pyo (2008, p. 3) argues that "fixed assets were evaluated at 'planned prices' and were deducted by only physical deterioration without accommodating the reduction of asset values due to obsolescence; they tend to be overestimated because the replacement value after the transition period becomes much higher due to the usual hyper-inflation during transition to market economies".

resources, i.e. a negative output gap. While the cyclical effect may capture the temporary movements of the economy, the structural output gap reflects the more enduring, embedded imbalances in these economies (skill mismatches, chronic current account deficits, high dependence on imports, high and structural unemployment, rigidities in the market, etc.) that are present regardless of the cyclical position in the economy.

The persistence of the negative output gap, associated with structural changes in early transition may also have been associated with more firms exiting the market than new firms entering, leading to a reduction in potential output. In addition, if actual output is smaller than the potential output, then large structural changes would lead to a persistent negative output gap. So, the question that arises is, 'why might actual output fall more than the potential output'? This is because potential output is technically determined by the factors of production. However, the major collapse of firms had a contagion effect on actual output, through the balance sheet effects (i.e. a negative multiplier).

Furthermore, due to a change in regime (i.e. centrally planned to market-oriented economy), the sectors that were previously vital in these economies (e.g. agriculture) gradually contracted and economic activity shifted towards other sectors. Therefore, previous labour skills were no longer required, generating a skill mismatch in the market. The skill mismatch, together with relatively high reservation wage and other market rigidities meant that the NAIRU was extremely high in most transition economies, whereas the actual unemployment rate could be much smaller. In addition, given the widespread obsolescence of the capital stock (especially in the early years of transition), even where significant physical capital was present, its value was fairly low. Hence, given the high NAIRU and obsolete capital do not directly enter into the calculation of the potential output, this leads to the conclusion that the potential output in these countries was low relative to the factors of production apparently available, hence the output gap was relatively small, which is reflected in the output gaps estimated and reported in sections 3.4 – 3.6 below. Therefore, the output gap in transition economies may to a large extent reflect the structural changes and to a smaller extent the cyclical component.

Thus, what is typically seen as an easy concept to define and apply, the output gap, in the transition context becomes rather complex and relatively difficult to define and measure. Hence, fitting a mean-trend onto the actual data, with say the Hodrick-Prescott filter, whereby the cycle would be below the actual output as much above, may not be appropriate in the transition context. Given the structural problems and the persistence of underutilised resources, other methods that would allow fitting a different mean are needed. This is why, as discussed in section 2.4.4, we decided to use the Unobserved Components method operationalized by the Kalman filter, as a method that accounts for structural changes as well as allowing fitting of the cycle reflecting persistent underutilised resources, i.e. negative output gap. A typical case would be the business cycle in Kosovo (see figure 3.6) where for most of the time observed actual output is below its potential, which reflects the economic characteristics of this country. On the other hand, in more developed European transition economies, where major structural changes have been largely completed, such as the Czech Republic (see figure 3.2), the differences with the developed countries have become smaller, thus the behaviour of the output gap reflects a more symmetric (developed country-wise) business cycle.

After discussing the conceptual definitions of potential output and the output gap, as well as their relevance in the transition economies context, we need to move on to the empirical modelling of these two indicators. Since, as mentioned earlier, potential output and the output gap are not directly observable (measured) they need to be estimated. Therefore, a full discussion of the most popular estimation techniques, as well as considering which is the most appropriate one for the transition context will follow in the next sections of the chapter.

2. 4. How are potential output and the output gap estimated?

As discussed in section 2.3, business cycle theory for transition economies is relatively underdeveloped. Furthermore, the lack of appropriate theoretical models (i.e. multivariate models, see section 2.5) specifically addressing the characteristics of transition economies, impelled some of previous studies to estimate potential output and the output gap using statistical approaches only (univariate techniques, see section 2.4.1), (Benes et al., 2005; Darvas an Vadas, 2006) or modify the mainstream theories (e.g. Phillips curve) to adapt the transition context (Gordon, 1997; Tsalinski, 2007; Bokan and Ravnik, 2012). Therefore, the lack of coherence between theory and empirical analysis led us to pay a considerable attention to estimation techniques and their appropriateness for transition economies.

Potential output is an unobservable variable, which cannot be found in any of the publicly available databases: it needs to be estimated independently by researchers. The importance of these estimates, as discussed in the previous section, encouraged economists and statisticians to develop many estimation techniques. However, in practice, different methods provide inconsistent results which pose problems in correctly identifying the cycle and in formulating macroeconomic policies (Canova, 1998; Cerra and Saxena, 2003; Benes et al., 2010; Biggs and Mayer, 2010). Thus, it has been argued, no matter how many methods one will use to estimate the output gap for robustness checks, there is no one most appropriate or correct approach and the generated estimates will be substantially uncertain (Fuentes et al., 2007; Magud and Medina, 2011). Hence, the unobserved nature of potential output and, thus of the output gap, together with the dispersion of the estimated results using different methodologies, makes their estimation inheritably difficult (Antoničova and Huček, 2005). Thus, as stated in the European Commission Report on Potential Output and the Output Gap (2001, p. 1), '[potential output and output gap] cannot be easily embedded in a robust and unquestionable quantitative indicators'. When estimating unobserved variables, such as potential output and output gap, two important questions arise: a) which method do we use for estimating unobserved variables and b) what observable variables do we consider to extract the unobserved variables from? The researcher can choose between a wide range of methodologies depending on the aim of the study, the appropriateness of the method and the data requirements. All the estimation methods for unobserved variables emanate from time-series models and can be classified as univariate or multivariate.[23] In order to understand the nature of the unobserved variables (potential output an output gap), we present and compare the advantages and limitations of the most commonly used univariate and multivariate techniques in the next section.

2.4.1. Univariate techniques

Most of the literature estimates potential output and the output gap through univariate techniques. The subject of this approach is mainly a single variable, in our case the actual output. These techniques generate values based on a purely mathematical/statistical basis, without considering the economic relationships to extract the unobserved variables, meaning that these techniques may be less accurate than multivariate ones. However, their simplicity makes them useful and quite popular among researchers (Anderton et al., 2014). In order to be able to describe and understand how unobserved variables (potential output and output gap) are generated, three methods will be discussed in this section: the linear time trend, the

[23]A diagram with 15 methods to estimate the output gap classified in four sub-groups is presented in Figure B2.1 in Appendix 2 of this thesis.

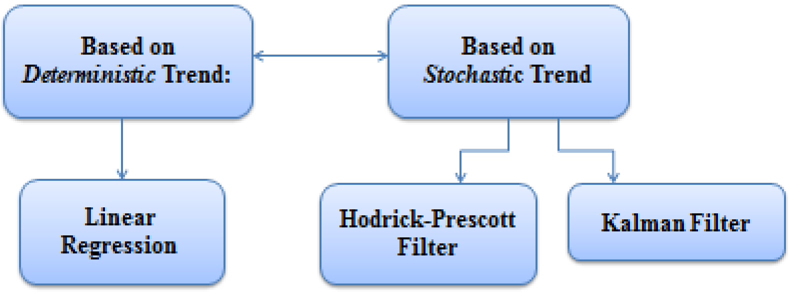
univariate Unobserved Components (UC), and the Hodrick-Prescott (HP) filter. For the purpose of exposition, univariate approaches will be discussed in depth, in order to clarify all their assumptions and estimation principles. Their multivariate extensions will then be presented more briefly, building on their respective univariate foundations.

The literature on univariate techniques has a long history in economic literature (Mitchell, 1946, Beveridge and Nelson, 1981, among others) and these are still widely used (Anderton et al., 2014). Almost all of the methods used (univariate and multivariate methods) to estimate potential output rely on detrending techniques, that is decomposing (filtering) the observed macroeconomic variable (actual output) into unobserved variables (components), that are trend, cycle and the corresponding unobservable parameters, which need to be estimated. The filtering procedure requires extracting lots of unobserved parameters which are usually included in general noise (e.g. the variance of trend and cycle). When using the univariate techniques, the ambitious task is to extract many unobserved variables and parameters from the noise using only one observable variable (in our case the real GDP). Filtering is a well-known concept when estimating unobserved variables or in forecasting, which means cleaning the data from the noise in order to identify the 'true' behaviour of the data over time. The noise concept in economics represents random departures from some 'signal' and some randomness surrounding it for no apparent reason. If there is more information left in the residual other than the randomness then that part of the process is not noise but rather some more information associated with the trend. Statistically or mathematically, this means removing (detecting) the trend from the cycle in a time series. In other words, detrending techniques decompose the actual output, which is an observable variable, into to the sum of the two unobserved variables, respectively the trend and the cycle:

$$y_t = g_t + c_t \tag{2.1}$$

where, y_t in our model stands for the actual output, the trend (potential output) g_t is usually used as an indication of the tendencies of the data over the long-run, whereas the cycle (output gap) c_t indicates the shorter run fluctuations of the time series. The decomposition procedure is a useful operation, since makes it easier to describe the behaviour of the observed actual output or other macroeconomic time series. We will elaborate two concepts of the trend, as presented in figure 2.1.

Figure 2.1. Approaches to estimating unobserved components based on different conceptual frameworks of the trend



In effect, linear regression, while not usually described as a filtering process, does act as a filter; from the observed data is filtering out an otherwise unobserved trend, and the difference between the trend and the actual data is the otherwise unobserved cycle. This is analogous to filtering out a signal, thereby isolating what is left out as a noise.[24]

2.4.2. The linear time trend

The linear time trend approach represents one of the oldest and most simple techniques for estimating potential output, which was widely used in the 1970s (Lucas, 1973, Barro, 1978). In this representation, macroeconomic time series were assumed to exhibit a relatively constant growing tendency (trend) over long periods, which can be modelled as a linear time trend by using the natural logarithm of the actual output. The mean reversion of the cyclical component on the other side is, by construction, purely transitory in nature, as it may fluctuate downward or upward over time; however, it cannot depart from the trend indefinitely. In this framework, no matter how far into the future periods we go, the deterministic trend will dominate the model. Thus the growth path of the time series will be quite predictable even in the longer run, regardless of the long-run economic performance. In the linear time trend approach, the observed actual output is decomposed into otherwise unobservable components:

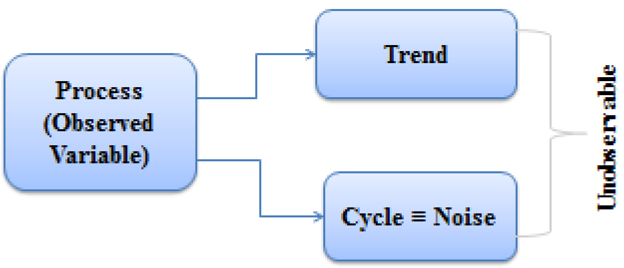
[24]The terminology of signal and noise originate from electrical and telecommunication engineering. Breaking up the process into Signal and Noise in engineering is analogous to the Trend and Cycle decomposition in the linear time trend representation in economics.

a linear long-run growth rate (trend); and a transitory component (cycle). This approach is set out in equation (2.2):

$$y_t = \alpha + \beta T + \varepsilon_t , \quad \varepsilon_t \sim iid(0, \sigma^2) \qquad (2.2)$$

where t = 1,…..n, represents the time component, y_t is the actual output, α is the time invariant constant term, T is the trend, whereas the cycle (output gap) is modelled through the residuals ε_t which are assumed to be stationary white noise, i.e. no trend in the residuals (Figure 2.2).[25]

Figure 2.2. Decomposition approach in the linear time trend



The residuals are assumed to be individually and independently distributed (henceforth ***iid***) with mean zero and variance σ^2.[26] The coefficient β represents the slope of the trend, which is assumed to be constant and stable over time (no structural breaks), defining the linear representation of the time series. So here, once we estimate α and β, given its linear representation, we have both the initial level (α) and the slope (β) and can predict where the trend is going in the long-run, meaning that the trend is completely deterministic over time.

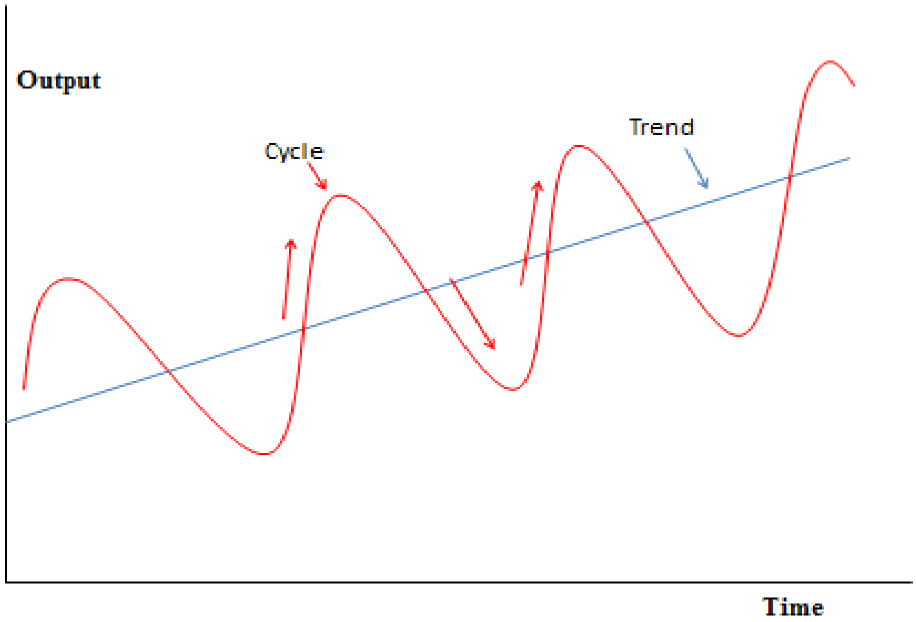
During the 1970s, the long-run trend was typically presented as a deterministic evolution path of the times series, where 'innovations' and 'surprises' were purely cyclical or stationary and could not impact the strong tendency of the trend (Figure 2.3). On the other hand, the output

[25] The distinction of the Trend and Cycle in the linear time *trend* representation (economic theory) is analogous to the signal and noise in engineering.

[26] The complex analysis of economic fluctuations such as *noise*, has led economists to consider these fluctuations as *iid* events (Bevilacqua and Zon, 2001).

gap, which represents the difference between potential output and actual output, was modelled by the whole residuals in the model, leading to an overestimation of the cycle, meaning that the cycle contains the residual part and may be prolonged or have a larger amplitude than the actual cycle. Further, the linear time trend framework assumed that all the variations ('surprises') in the model are captured by the residuals and that they will not impact the strong tendency of the deterministic trend, because the actual path will never go far from its trend.

Figure 2.3. Linear time trend



Even though the linear time trend method enjoys the virtue of simplicity, it was criticised on the grounds of imposing too many restrictions on the model, since an intercept, slope and a stable trend may not be sufficient to explain variations in output. In particular, the linear time trend actually failed to account for key external shocks that happened during the 1970s when the oil shocks disturbed the data causing autocorrelation and non-normality in the residuals. The failure of this simple representation questioned the model's basic assumption that the output data is trend-stationary. If this assumption is false then the time-trend regression models are misspecified. Therefore, later on, contrary to the constant growth rate of the trend as proclaimed in the deterministic view, a new set of approaches that allowed the growth rate of potential output to vary started to develop. Some of the most popular of these approaches that model the trend as a random walk with drift will be explored below.

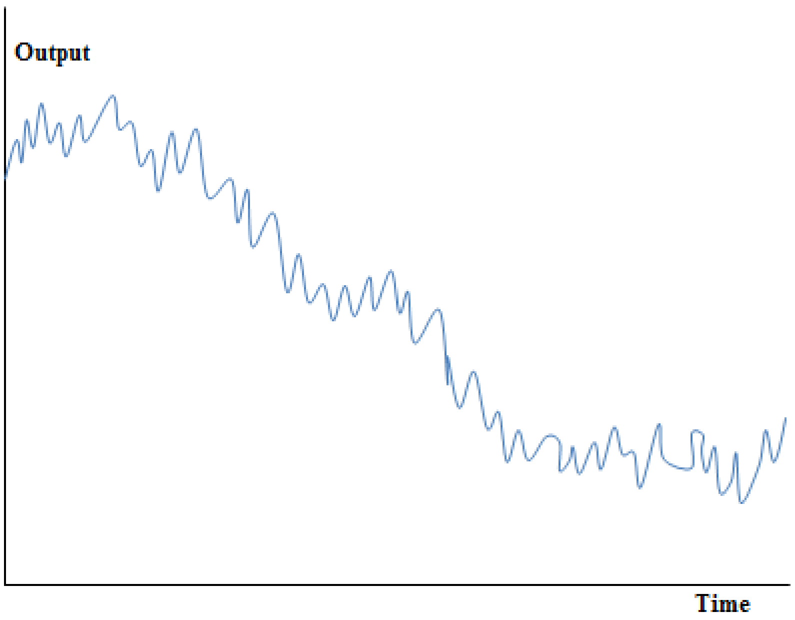
This approach to unobserved components separates out a purely deterministic trend and a cycle which is stationary process around a trend. Since the 1970s, both the concept of the trend and the concept of the cycle have been subject to criticism and consequent reformulation. In brief, the concept of the trend has been enriched by incorporating unit root processes and the cycle by being modelled as an autoregressive process. The next two sections deal, in turn with each of these. We then proceed to show how these developed concepts of both the trend and the cycle have been incorporated into modern unobserved components methodology, in particular, the Kalman filter.

2.4.3. Modelling Trend as a Random Walk with Drift

The limitations of the linear time trend models led authors like Beveridge and Nelson (1981), Nelson and Plosser (1982), and later Clark (1987) to enquire into the possible misspecification of the magnitude and persistence of business cycles in linear models.

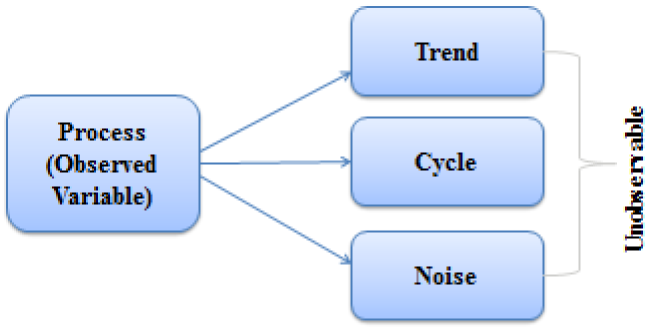
Nelson and Plosser (1982) and Watson (1986) challenged the linear trend concept, by arguing that the growth trend of economic time series is stochastic, making the basic assumption that the output data are trend stationary and that the cycle is modelled by the whole residuals in the model debatable (Figure 2.4).

Figure 2.4. Random walk time series



If this is the case, then modelling macroeconomic time series with a linear time trend will yield statistically misspecified trend and residuals. Indeed, they argued that macroeconomic series, like output, represent stochastic processes where the trend (potential output) is typically modelled as a random walk with a drift term, and that the cycle (output gap) should also be filtered out /separated from the residuals (Figure 2.5).[27] After challenging the linear time trend concept, a new generation of techniques drew upon the random walk and the unit root theory, as part of the new approach.[28]

Figure 2.5. Decomposition approach in the random walk process



Beveridge and Nelson (1981) were among the first to present a procedure for decomposing non-stationary time series into permanent and transitory shocks, where both can be treated as stochastic. They argued that each ARIMA process has a stochastic trend and can be decomposed into two unobserved elements: a stationary series and a pure random walk. Since most of the macroeconomic time series tend to be integrated of order one-I(1) (Nelson and Plosser, 1982; Harvey and Jaeger, 1993; Cochrane, 2005), they can be modelled as an autoregressive-integrated-moving-average process. Further, Stock and Watson (1988) found

[27] ‘A random variable is a function whose values depend on the outcome of a chance event. The values of a random walk may be any convenient mathematical entities; real or complex numbers, vectors, etc…..Sums, products and functions of random variables are also random variables’ Kalman (1960, p. 45). Noise in random walk processes can be described as normal distribution, the part that is not random is part of the signal.

[28] A unit root process is one that is integrated of order one, meaning that the process is non-stationary. In this case, the first differencing of the process will usually yield stationary time series. The simplest example of the unit root process is the random walk (Hamilton, 1994). Unit root and random walk processes represent two common stochastic trends.

that the actual output can be decomposed into two unobservable components, a non-stationary and a stationary component. The stationary element being the cycle (output gap), which is defined as points in time fluctuating around a fixed mean (with stable autocovariance), and the random walk with drift representing the trend path (potential output) of the series. They also argue that most of the variation in the actual output comes from the variation in trend, whereas the variations that come from the cycle component are smaller and rather noisy. (This decomposition approach represented the base framework for another approach, namely Unobserved Components approach, which will be elaborated in the next section). Formally, the trend component conceptually modelled as a random walk with drift, and a stationary cyclical component modelled as an autoregressive process follow the form:

$$y_t = g_t + c_t \qquad (2.1)$$

$$g_t = \underbrace{\mu_g}_{\substack{\text{deterministic}\\ \text{component}}} + g_{t-1} + \underbrace{\varepsilon_t}_{\substack{\text{random}\\ \text{component}}}, \qquad \varepsilon_t \sim iid(0, \sigma^2_\varepsilon) \qquad (2.3)$$

$$\phi(L)c_t = \upsilon_t, \qquad \upsilon_t \sim iid(0, \sigma^2_\upsilon) \qquad (2.4)$$

In the (2.3) equation, t represents the time index $t = 1,\ldots..n$, the trend g_t is modelled as dynamic linear stochastic process which depends on disturbances and is a random walk with drift process. The drift is represented by the constant μ_g. The error term ε_t is assumed to be white noise, with zero mean and variance σ^2_ε. In this framework, the trend (potential output) is represented as a random walk with drift, which will increase in each period by a constant, i.e. the drift:

in period t = 1, $\Rightarrow g_1 = \mu_g + g_0 + \varepsilon_1$ (2.5)

in period t = 2, $\Rightarrow g_2 = \mu_g + g_1 + \varepsilon_2$ (2.6)

$\Rightarrow g_2 = \mu_g + (\mu_g + g_0 + \varepsilon_1) + \varepsilon_2$ (2.6a)

in period t = 3, $\Rightarrow g_3 = \mu_g + g_2 + \varepsilon_3,$ (2.7)

$\Rightarrow g_3 = \mu_g + [\mu_g + (\mu_g + g_0 + \varepsilon_1) + \varepsilon_2] + \varepsilon_3$ (2.7a) leading to the generalised form, where all the past changes are accumulated:

$g_t = t\mu_g + g_0 + \sum_{t=1}^{n} \varepsilon_t$ (2.8)

The equation (2.4) approximates the stationary cyclical component c_t through an autoregressive process (AR). Here, the ϕ coefficient stands for the autoregressive coefficients, the (L) is the lag operator and the υ_t is the *iid* error term.

Contrary to what happens in the case of deterministic trends where shocks have only a transitory impact, the random disturbances ε_t, or shocks, cumulate from the start of the process, without being absorbed in the future, i.e. the impact of a shock does not go away (shocks have a permanent influence), (Cochrane, 1988). The drift term μ_g, multiplied by the number of periods *t*, accumulates in each additional period, this way giving direction to the series and giving weight to the deterministic part of the process. However, given the random nature of the process and the corresponding unpredictable component over the long-run, the series may deviate from the average in each period by an unforeseeable amount. This is because, in the random walk with drift process, the series may move far away from the mean, and the farther it goes into the future, the less predictable it becomes, so the random component of the series dominates the process. The number of error terms in the trend component increases in each period, e.g. in the third period, the trend component has three error terms, in the n_{th} period, the trend has *n* error terms, and so on. This makes the trend component a more fluctuating process, hence a more complex one to estimate and forecast, relative to the case with a deterministic trend. Nelson and Plosser (1982) and Stock and Watson (1988) have shown that the 'varying trend' is present in most of the macroeconomic time series that follow a stochastic process, which could be explained by the presence of random shocks with persistent effects. The persistence of the error terms over time may naturally lead to the presence of the unit root in the stochastic trend. In the unit root processes the trend is non-stationary and follows a random walk (often with a drift):

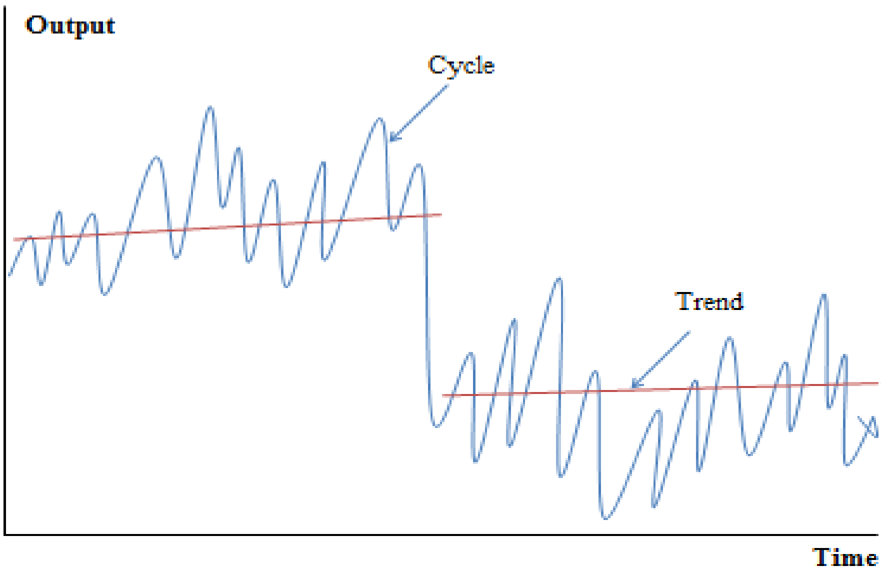
$g_t = \rho g_{t-1} + \varepsilon_t$ with $\varepsilon_t \sim iid(0, \sigma^2_\varepsilon)$ and $\rho = 1$ (2.3a)

so, if ρ is equal to 1 or close to 1, then each error ε_t has persistent effects on the trend g_t, since each fluctuation will not be absorbed in the future. In this case, the trend is generated by past and present noise ε_t.[29] In line with this, Nelson and Kang (1981), Nelson and Plosser (1982) and Perron (1986) empirically demonstrate that many macroeconomic time series follow a stochastic process (starting from GNP, Consumption or Stock Prices), i.e. are non-stationary in levels and are integrated of order one-I(1)[30]. Nelson and Plosser (1982) test the difference-stationary hypothesis against the trend-stationary hypothesis, and find that macroeconomic time series are correlated with a unit root relationship. They suggest to take the first-difference (of actual output, trend and cycle) in order to take out the non-stationarity (growth) and make the *trend* a random walk process with a drift, rather than a 'straight line', and the cyclical component (output gap) a stationary process.[31] After first-differencing the series, they find that most of the variation in the series should be allocated to the non-stationary trend, with much less variation allocated to the cyclical component. However, Nelson and Plosser (1982) accept that customary testing (Dickey and Fuller test) reject the null in favour of stationarity far too often, in both cases, so the diagnostics are usually weak.[32]

Later, Perron (1989) showed that, what may seem as a unit root process, e.g. a random walk with a drift process, may in fact represent a deterministic trend with breaks and shifts, at least for some of the macroeconomic time series (Figure 2.6).[33] He showed that when fluctuations are stationary along a broken trend, the Dickey-Fuller test is not able to reject the unit root hypothesis. Perron (1989) reintroduces the older idea of the linear time trend but develops the concept of the trend, as a possible compromise between the two earlier *trend conceptual* frameworks, random walks and linear deterministic trends.

[29] Since, noise is an *iid* and exogenous variable, in the presence of a unit root, the trend entirely depends on the error terms, which are of unknown nature (Bevilacqua and Zon, 2001).
[30] A process (e.g. trend) that has a unit root is also called integrated of order one, denoted as I(1); conversely a stationary process is an integrated of order zero, I(0).
[31] The non-stationarity can be eliminated by taking the first difference of the series, eventually leading to a stationary series with a constant mean and variance. However, the empirical work of Nelson and Plosser (1982) relies on the strong assumption that after the first differencing of the output, the autocovariance function of output is zero after one lag (Clark, 1987).
[32] Failure to account for the structural breaks in the series, i.e. change in the slope of the series, will provide weak ADF unit root test results.
[33] Perron (1989) developed a test that enables rejecting the unit root hypothesis in the time series with broken trends, after he arbitrarily assigned the date in which the structural break has occurred.

Figure 2.6. Broken linear trend



To sum up, the discovery that the trend component may not evolve along a linear ‘straight’ line, as assumed in classical econometrics, but rather as a random walk with drift process, represented a radical change in the conceptualisation of the trend, and has deeply influenced the direction of the empirical research in recent decades. However, though contributing to a shift in thinking how macroeconomic time series should be modelled by challenging the concept of the deterministic trend, authors like Beveridge and Nelson (1981) and Nelson and Plosser (1982) failed to explain why other functional forms, such as nonlinear or stochastic trends were not appropriate for modelling trend and cycle. Moreover, even after detrending or first differencing the data, which eventually leads to stationary series, these approaches could not predict events or shocks (called as ‘innovations’) and explain structural breaks in the series. Perron on the other side did not reject the idea of unit root process (random walks) *per se*, but he established the idea that testing the idea of unit root should be done conditional on testing with deterministic time trend with one or more structural breaks.[34] Further, Bevilacqua and Zon (2001, p. 14) argue that “the random walk hypothesis is not generally rejected by the unit root literature and it is the core of the business cycle theory”.

[34] Perron (1989) argued that over long periods, deterministic time trend do not go on forever, they are subject to breaks (e.g. the Great Depression in 1929, oil price shocks in 1970s). In line with this, Perron (1989) showed that a series which is stationary around a deterministic time trend but has undergone a permanent shift some time ago, will provide unit root results.

Overall, stochastic processes are more appropriate in handling the estimation of *stochastic trends* and for modelling the long-run uncertainty, however they do not do all the decomposition needed, being unable to estimate the shocks and shifts. The limitations of the previous models gave way to another method for detrending unobserved variables, namely the unobserved components method, which will be elaborated in the following section.

2.4.4. Unobserved Components Model (a.k.a. the State Space Model)

The unobserved components (UC) model was introduced into the economic literature by Harvey and Todd (1983), with Watson (1986) and Clark (1987) being among the first authors to implement it in empirical studies. These studies carried forward the conceptual revolution in the understanding and analysis of economic *trend*; of the trend into the UC approach. In contrast to the empirical study of Beveridge and Nelson (1981), the UC approach implies a smoother trend, but a rather high amplitude and persistent cycle. Similar to the linear time trend approach, decomposing the linear trend and cycle is also the main idea of the UC approach; even though this term was not used in the literature of 1970s. The main objective of the UC model also lies in decomposing and estimating unobserved components (variables) such potential output, output gap, natural rate of unemployment, etc. using all the information available from the observed variables (e.g. actual output, unemployment rate, inflation rate etc.); in other words, extracting the signal from the noise (signal processing technique). However, the UC approach is much richer and complex in structure, and models like least-squares or instrumental variables cannot be used to solve the model. Hence, the more advanced framework of the UC model enables the estimation of unobserved variables with less error. The UC method represents a flexible framework, where all components may be optional, that is, one can add seasonal effects into the equation in the case of non-seasonally adjusted data, or the trend may be first-order deterministic or second-order stochastic process and the cyclical component may be either stationary or stochastic process (Harvey and Trimbur, 2008). For example, Watson (1986) and Kuttner (1994) assume that the drift term is constant in contrast to other more general cases where the drift is also a time-varying component in a macroeconomic time series. This assumption may be more suitable for more steady-state economies such as the United States, however, in other countries such as European transition economies (see the discussion in section 2.3), where pronounced changes in the trend growth rate take place, this assumption may not be appropriate. Clark (1989) for example, assumes that the drift also follows a random walk:[35]

$$\mu_t = \mu_{t-1} + u_t \tag{2.9}$$

Todd and Harvey (1983) on the other hand, allow both the slope and the drift to be time varying. One of the simplest applied forms for economic data is the Watson (1986) model, where the potential output modelled by the trend is modelled as a stochastic process, represented by a random walk with drift process and the output gap denoted by the cyclical component is an autoregressive process of order two. The UC representation takes the following form:

$$y_t = g_t + c_t \tag{2.1}$$

where the trend is modelled as a random walk with drift:

$$g_t = \mu_g + g_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma^2_\varepsilon) \tag{2.3}$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \upsilon_t \qquad \upsilon_t \sim iid(0, \sigma^2_\upsilon) \tag{2.10}$$

In this framework, t represents the time index t = 1,…..n. The trend g_t is a random walk process. The error term ε_t is assumed to be individually and independently distributed over time, with mean zero and variance σ^2_ε. Thus, the disturbance is serially uncorrelated and independent of any other disturbances related to y_t in equation (2.1). Here, g_t and σ^2_ε are unobserved and thus need to be estimated. The drift is represented by the constant μ_g which stands for the average growth rate of potential output.

The cycle c_t, which by definition denotes the output gap, is an autoregressive process of order two AR(2), with autoregressive coefficients ϕ_1 and ϕ_2. The output gap is assumed to follow an autoregressive process of order two-AR(2); i.e. two lags and a zero unconditional mean (Gerlach and Smets, 1997). Authors like Nelson and Plosser (1982) and Clark (1987) decided that the output gap should follow an AR(2) process mainly based on empirical grounds, rather

[35] Kichian (1999) explains that the reason for including two different error terms is for expositional and technical convenience, since it allows the computer programming to be more tractable.

than having any explicit theoretical explanation.[36] The error term ε_t in (2.3) stands for permanent shocks whereas the error term υ_t in (2.10) stands for transitory shocks which are assumed to be serially uncorrelated and independently distributed from each other.[37] In contrast to the Beveridge and Nelson (1981) framework, where the trend also follows a random walk and the errors of the trend and the cycle may be correlated, in the UC framework the cycle and trend shocks are assumed to be uncorrelated over time. Thus, the Beveridge and Nelson (1981) framework is sometimes also called a form of the UC representation.[38] The assumption that the trend and cycle error terms are not correlated was firstly posed by Kalman (1960)[39], which was followed by Watson (1986), Clark (1987) and many of the more recent studies, which make this assumption by default (Kichian, 1999; Welsch and Bishop, 2006). In the UC approach, this assumption represents the main restriction in the UC representation. In random walk processes where transitory and permanent shocks have different impacts on the actual output, an output gap appears. Thus, a diffusion process of the two shocks is desirable, "since the economy is likely to remain in its production possibility frontier as adjustments unfold" (Darvas and Vadas, 2003, p. 43). Clark (1987) regards this assumption as false, given that even transitory shocks may have an impact on the permanent component (potential output). For example, an economic shock may reduce investment and hence, the capital stock via an accelerator mechanism, thus it may impact potential output also. However, the prior assumption that innovations or shocks affecting the potential output and the cycle are not correlated is necessary to make the model econometrically identifiable (Kuttner, 1994; Gerlach and Smets, 1997), which assumption enables the identification of the transitory shocks with and without impact on the permanent component. Allowing the σ^2_υ to be positive and σ^2_ε to be zero, the estimated trend will be relatively smooth (Harvey, 2005). Nevertheless, in our representation, both variances are allowed to vary.

---

36 Milionis (2003, p. 3) argues that, "in practice, it is difficult to test for strict stationarity. Instead, the so called 'weak' or 'wide sense stationarity' up to the second order can be testable".

37 However, only the 'new information' triggers a revision in the permanent component. The cyclical part is serially random by definition, and is proportional to the current innovation and therefore also to the current innovation in the permanent part (Beveridge and Nelson (1981, p.156).

38 Morely, Nelson and Zivot (2002, p. 8, 9) argue that "there is always at least one UC representation for any given ARIMA process". Also, comparing the Beveridge and Nelson (1981)-BN work with the UC framework, they argue that "the BN trend is the conditional expectation of the random walk component for any UC representation of an I(1) process".

39 Kalman (1960, p. 35) states that "…arbitrary random signals are represented (up to second order average statistical properties) as the output of a linear dynamic system excited by independent or uncorrelated random signals (''white noise'')…".

The UC method can be implemented using the Kalman filter, with the trend modelled as a random walk with drift and the cycle as an AR(2) process, once it is written in a state space notation (see below).[40] A state space representation of the UC approach means formulating all the components of interest, that in our case are potential output, output gap, parameters and error terms as a system of matrices and vectors. Afterwards, each component is separately modelled by an appropriate dynamic stochastic process, which usually depends on normally distributed disturbances (Koopman et al., 2008). This approach basically groups all the observed and unobserved variables in two equations, the so called *measurement* or observation equation and the *transition* or state equation, otherwise representing two algorithms in the Kalman filter (Commandeur and Koopman, 2007).[41] In our case, the measurement equation is formulated in equation (2.1), whereas the transition equations are equations (2.3) and (2.10). The measurement equation links the (1 x 1) vector of observed variables Y_t (actual output) with the (3 x 1) vector of unobserved variables X_t (potential output, output gap and lagged output gap) through the vector of exogenous variables D_t. The transition equation describes the dynamics of the state vector, where vector of exogenous variables enter into the equation. In the state space framework, the observed variables are known as ‘signal’ variables, whereas the unobserved variables are known as ‘state’ variables (Kalman, 1960). The error term is the vector of serially independent disturbance series. Firstly, we present signal and state variables as equations, and then transform them into a state space form of vectors and matrices, in a constant-parameter linear state space model (Watson model):

The measurement (observation) equation:

$$Y_t = ZX_t + DD_t + \varepsilon_t , \qquad \varepsilon_t \sim N(0, \sigma^2_\varepsilon) \tag{2.11}$$

where Y_t is the (1 x 1) vector of the observed variables, X_t is the (3 x 1) vector of the unobserved variables, Z is the (1 x 3) matrix of the coefficients of unobserved variables, and in the case where other exogenous variables are included (e.g. inflation in the Phillips curve or unemployment rate in NAIRU, in multivariate representations of the UC model), D will stand for the matrix of the coefficients of the exogenous variables and D_t is the vector of

[40] Welch and Bishop (2006) explain that the filter is very powerful in several aspects, since it can handle past, present and future states even when the nature of the system is unknown.

[41] Koopman et al., (1998) explain that in order to generate samples from the unconditional distributions implied by a statistical model in state space form, or to generate artificial datasets, we use the state space form as a recursive set of equations.

exogenous variables. In the univariate representation of the state space model, the exogenous variables are missing since we extract the unobserved potential output and output gap by using the actual output as the only known information. In the univariate form of the measurement equation, we do not include exogenous variables (e.g. inflation rate, unemployment rate, etc.), so this term is equal to zero, and the restricted Watson (1986) form will be:[42]

$$Y_t = ZX_t \tag{2.12}$$

As mentioned earlier, in the UC model, the framework should be written in a state space form, that is, as a system of vectors and matrices. In the state space form, the expression in (2.12) can equivalently be written as:

$$Y_t = [y_t],$$

$$Z = [1 \;\; 1 \;\; 0], \text{ and}$$

$$X_t = \begin{bmatrix} g_t \\ c_t \\ c_{t-1} \end{bmatrix}.$$

Writing these expressions as a system of vectors and matrices will give us the explicit form of the measurement equation:

$$[y_t] = [1 \;\; 1 \;\; 0] * \begin{bmatrix} g_t \\ c_t \\ c_{t-1} \end{bmatrix} = g_t + c_t \tag{2.13}$$

The measurement equation relates the observation y_t to the state vector X_t through the signal ZX_t. Here the algorithm of the Kalman filter does the prediction of the unobserved components.

The second equation is the *transition equation (state)* of the UC model:

$$X_t = BX_{t-1} + AZ_t + \upsilon_t \,; \qquad \upsilon_t \sim N(0, \sigma^2_\upsilon) \tag{2.14}$$

[42] In the univariate approaches, unobserved components, such as the trend and the cycle, are filtered out of some single series (e.g. actual output). Conversely, in multivariate approaches, this filtering is accomplished conditional on values of other variables, which are treated as exogenous.

Where B is the (4 x 4) matrix of coefficients of unobserved variables, A is the matrix of coefficients of exogenous variables and the Z_t is the vector of exogenous variables.

In the state space form, again, since we are presenting the univariate framework with no exogenous variables, the term AZ_t is omitted. So, presenting the transition equation in the state space form, where:

$$X_t = \begin{bmatrix} g_t \\ 1 \\ c_t \\ c_{t-1} \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & \mu_g & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & \phi_2 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$X_{t-1} = \begin{bmatrix} g_{t-1} \\ 1 \\ c_{t-1} \\ c_{t-2} \end{bmatrix}, \text{ and}$$

$$\varepsilon_t \ and \ \upsilon_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \upsilon_t \\ 0 \end{bmatrix}.$$

Here, the lagged signal (4 x 1) X_{t-1} indicates that the transition algorithm is a recursive process (see section 2.4.6). The combinations of all these vectors and matrices will give us the explicit state space form of the transition equation:

$$\begin{bmatrix} g_t \\ 1 \\ c_t \\ c_{t-1} \end{bmatrix} = \begin{bmatrix} 1 & \mu_g & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & \phi_2 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} g_{t-1} \\ 1 \\ c_{t-1} \\ c_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ \upsilon_t \\ 0 \end{bmatrix} \tag{2.15}$$

Generating:

$$g_t = g_{t-1} + \mu_g + \varepsilon_t \quad \text{(a random walk with drift)} \tag{2.3}$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \upsilon_t \quad \text{(an AR(2) process)} \tag{2.10}$$

The vectors of error terms, following the standard assumptions in the state space methodology, are assumed to be normally and independently distributed, uncorrelated with each other, with mean zero and variances σ^2_ε and σ^2_υ.[43] The transition equation is actually the recursive process of updating the estimated data. It formulates the dynamic process of the unobserved components in the model.

In this representation of the UC model, the vector of observed variables Y_t contains only actual output, whereas the vector of unobserved variables contains potential output (g_t), the cyclical component (c_t) and the first lag of the cyclical component (c_{t-1}). Here, the initial parameters μ_g, ϕ_1, ϕ_2 are treated as unknown, whereas the remaining part needs to derive its statistical properties. Thus, in a state space representation we are faced with two groups of unknown parameters that need to be estimated, for a given model specification. The first group represents the trend (g_g), the cycle (c_g) and the drift (μ_g), whereas the second group are the initial parameters, ϕ_1, ϕ_2 and the error variances (σ^2_ε and σ^2_υ). While there are thus seven unknown components that need to be estimated, there is only one piece of information from which to extract them, which is actual output (real GDP).

The UC modelling in the Watson (1986) form represents a simplified (restricted) version, as it does not include exogenous shocks, dummy variables or any additional explanatory variables such as inflation rate or unemployment rate variables that apply to multivariate techniques. We use it here to provide an intuitive and general explanation of the formulated state space form. Different authors modify and augment the state space notation by adding exogenous or dummy variables, depending on the aim of their study. For example, Kichian (1999) argues that macroeconomic series characterized with high frequencies usually exhibit thick tailed distributions, which may be explained by the conditional heteroskedastic variances of these time series. This may be the case in models (countries) that have experienced several changes/switches in regimes: the period when countries switched from planned economy to market oriented economy, or countries that have switched their monetary regimes several times during the transition periods, consequently having a changing mean and variance of inflation, etc.

[43] Since the mean is assumed to be zero and the error terms σ^2_ε and σ^2_υ are assumed to be mutually uncorrelated, then the correlation and the covariance are the same (Wooldridge, 2005).

Once written in a state space form, the UC model can be operationalized using the Kalman filter. The UC approach, written in a state space form, represents the structure of the Kalman filter. A Kalman filter is implemented conditional on the initial parameter values (μ_g, ϕ_1, ϕ_2, σ^2_ε and σ^2_v), which 'design' the Kalman filter on modelling the desired components of our model. The initial parameter values need to be assumed or calibrated before implementing the Kalman filter. The parameters of the UC model can then be estimated using the maximum likelihood function. The criteria for determining the value of the initial parameters will be such that they will minimize the difference of the mean squared error E{ $|y_t - \hat{y}_t|^2$}, (further discussion and estimation of the initial parameters is postponed until chapter 3).

2.4.5. The Kalman filter

The Kalman filter was first developed by Rudolph E. Kalman in 1960 as a system for modelling discrete linear data with random processes. The Kalman filter represents a recursive process that generates *optimal estimates* from indirect, inaccurate and noisy observations.[44] In the case of stochastic processes which involve noisy data, the Kalman filter 'cleans up the noise' in order to extract the true value of the desired variable together with its spread (behaviour). If all noise is random (Gaussian), the Kalman filter minimises the mean square error of the estimated parameters, in order to generate the *optimal estimator* (Kleeman, 1995).

In the case of discrete linear data modelling, the filter is always designed as an algorithm.[45] In our case, the filter is designed from two algorithms that are measurement and transition equations. Algorithms serve as a guide to predict the unobserved components, where entering a set of observed and otherwise assumed information (y_t, μ_g, ϕ_1, ϕ_2, σ^2_ε and σ^2_v) in the system will generate an outcome (g_t and c_t).[46] These parameters, together with other observed variables in the case of multivariate UC model, feed in the system and basically design ('train') the Kalman filter in how to filter or predict the future values of the unobserved variables. The parameters can be changed depending on the needs of the model; however the

[44] It is a recursive process in the sense that new measurements can be processed as they arrive.

[45] Hamilton (1994) argues that numerical algorithms are usually much better behaved if an 'intelligent' initial guess of parameters is used.

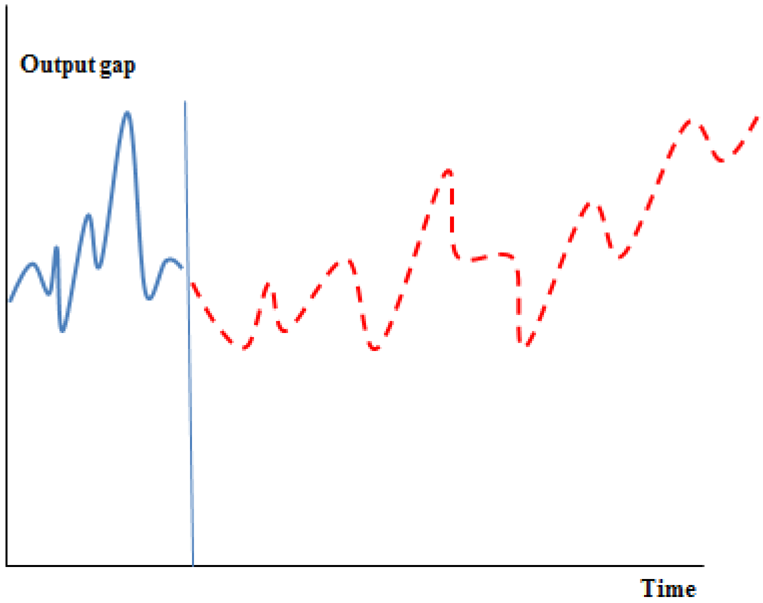
[46] The ratio of $q = \sigma^2_\varepsilon / \sigma^2_v$ represents the signal-noise ratio, which determines how observations should be weighted for prediction and signal extraction. The higher is q, the more past observations are discounted in forecasting the future. Similarly a smaller q means that the closest observations receive a bigger weight when signal extraction is carried out. Further, a higher q means that most of the variation in the series is allocated to the trend component, whereas a smaller q means that most of the variations are allocated to the cyclical component.

number of the parameters remains fixed. The Kalman filter is highly sensitive to initial parameter values, which in effect will guide the way to predict unobserved components, thus the starting judgement for them is particularly important. For example, if the series is completely driven by structural shifts, the initial parameter values should be assumed differently compared to the case where the series is assumed to be driven by smooth fluctuations. In transition economies, a higher weight should be assigned to the trend variance σ^2_ε, as compared to, say, to the pre-financial crisis in the developed countries, where the trend is pretty smooth.

The lagged terms in the transition equation indicate that the transition algorithm in the Kalman filter will be a recursive process. A recursive process in the discrete data modelling means that each additional observation entered into the algorithm will predict each future values (one-step ahead predicted errors) based on past values at time t, then the following predicted values ($t + 1, t + 2, \ldots t + n$) will be updated (estimated) based on all past values plus the new estimated predicted value and so on. This process is illustrated in Figure 2.7, where the blue line represents the entrant data (the input in the algorithm of the Kalman filter), whereas starting from the dashed red line, each discrete value is being predicted (output of the filter), one by one, creating the behaviour of the series of interest (output gap). The recursive process runs into every term in the measurement and transition equation, similarly for trend, for cycle and for each parameter in succession.[47] After much iteration, the algorithm will eventually converge into an optimal estimator. Estimating the unobserved components one by one at each point in time, 'allows the data to speak', meaning that the algorithm is able to identify itself the structural breaks in the sample, without having to impose external restrictions.

[47] Welch and Bishop (2006, p. 5) explain that the recursive nature of the Kalman filter represents one of its most appealing features, since it makes the practical implementations more feasible, compared to other filters (e.g. Wiener filter) which are designed to operate on all of the data directly for each estimate.

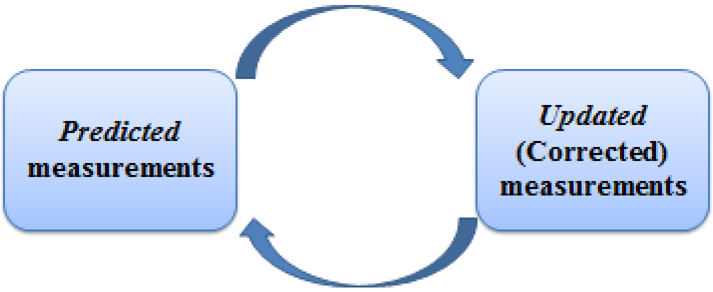
Figure 2.7. The recursive process in Kalman filter



The estimation procedure using the Kalman filter is as follows. The first step is to write the model in a state space form, as denoted in equations (2.11) to (2.15). Second, it requires setting up the initial parameter values of the model (μ_g, ϕ_1 and ϕ_2), which in our case, are time invariant and characterize the optimal linear state of estimation, i.e. define the model as filtering or prediction. Also, the filter requires setting the starting values of the unobserved variables (g_t, c_t) and the variances of the unobserved variables and their co-variance matrix cov(σ^2_ε, σ^2_v)) which in effect will define the initial *state of the model*. The starting values of the unobserved components (g_t, c_t) and other parameters are generated from other methods, such as OLS, calibrated or assumed, thus not generated from the filter itself. In the third step, the transition algorithm (equation (2.14)) generates the current state variables, conditional on the initial parameter values. These are the *predicted* values of the unobserved variables (potential output and output gap), which still contain some noise (error) and uncertainties in them. This stage is also known as the *prediction step* or *a priori* estimate for time t.

In the fourth step, also called the *updating (correcting) step*, the 'cleaning' process continues, where the predicted values of unobserved variables are cast into the measurement algorithm (equation 2.12)), to generate the predicted values for the observed variables. The predicting-correcting process in the Kalman filter represents the feedback control process (Figure 2.8).

Figure 2.8. The ongoing *predict-correct* process in the Kalman filter



Source: Welch and Bishop (2006).

After each predicted and updated measurement pair, the process is repeated with the previous estimates used to project or predict new estimates, until the parameter estimates converge (Commandeur and Koopman, 2007).[48] The difference between the observed actual output and the predicted output $E\{|y_t - \hat{y}_t|\}$ is called the measurement innovation or the residual, which quantifies the lack of the accuracy of the estimated unobserved estimates in predicting the observed values at time t (Welch and Bishop, 2006). Otherwise, the measurement innovation looks like (Hamilton, 1994):

$$E\{|y_t - \hat{y}_t|\} = min\{(\sigma^2_\varepsilon - \widehat{\sigma^2_\varepsilon}) + min\,(\sigma^2_\upsilon - \widehat{\sigma^2_\upsilon})\} \quad (2.16)$$

Simultaneously, the Kalman filter also calculates the prediction error and the variance terms (σ^2_ε and σ^2_υ). If the measurement noise vector components (ε_t, and υ_t) are uncorrelated, then the state update can be carried out one measurement at a time. Here, in a recursive manner, the Kalman algorithm updates the predicted values by giving a higher weighted average to the estimates with higher certainty and vice versa (as measured by q, see note 18). The unobserved variables will be estimated using maximum likelihood function (de Brouwer, 1998):

[48] Given a set of particular data with a given set of parameters, the maximum likelihood will maximise its function. However, for a given set of information, sometimes the maximum likelihood cannot maximise the maximum likelihood function. One possibility to improve the maximum likelihood function would be to add additional information in it.

$$log\Lambda = -\frac{NS}{2}\, log\, 2\pi - \frac{1}{2}\sum_{t=1}^{S} log\,|F|_t - \frac{1}{2}\sum_{t=1}^{S} \nu_t'\, F_\nu^{-1} \quad (2.17)$$[49]

"….which reflects how likely would have been to have observed the data (potential output and output gap) if the initial parameter values were the true values…" (Hamilton, 1994, p. 3055).

The value of the maximum likelihood function is maximised by minimizing the prediction errors (ε_t, and v_t) and their variances (σ^2_ε and σ^2_v). If the likelihood function is maximised, conditional on initial parameter values and initial values of g_t and c_t (Hamilton, 1994):

$$\log\,(y_t, y_{t-1}, \ldots, y_1 |\, X_t, X_{t-1}, \ldots, X_1; \mu_g, \phi_1, \phi_2, \sigma^2_\varepsilon, \sigma^2_\upsilon) \quad (2.18)$$

then the estimated unobserved (g_t and c_t) components and the parameters (μ_g, ϕ_1, ϕ_2, σ^2_ε and σ^2_v) are maximum likelihood estimates; if not, then the algorithm will continue a set of iterations and recursive processes by iterating on measurement and transition equations, using new initial parameter values, until it generates the optimal error estimates (Hamilton, 1994).[50] The maximum likelihood function, conditional of the initial parameter values, maximises the log-likelihood function value, by minimising the mean squared error, in order to generate the optimal estimates. The initial values of the parameters are the values that make equation (2.16) as large as possible; nevertheless, any change in the parameter values would imply a different probability distribution. The Kalman filter may be considered as one of the best linear estimators that outperforms other models, (e.g. ARIMA models), (Koopman et al., 2008). Thus, whenever modelling involves considerable noise, such is the case with unobservable components like potential output and output gap, it is usually suggested to use a Kalman filter (Kleeman, 1995).

Overall, the UC method written in the state space notation has several advantages relative to other univariate methods. First, the UC method written in state space notation offers enormous flexibility when dealing with irregular data, such as missing observations and observations with mixed frequencies such as monthly and quarterly (Harvey, 2005). In the

[49] Where, N-number of observed variables, S-sample size, ν- prediction error matrix, F-mean square error matrix for the prediction errors.

[50] Hamilton (1994) explains that the mean squared errors are not a function of the data y_t and they can be evaluated without calculating the forecast for the unobserved components.

case of missing values, using the Kalman filter is practically convenient, since the filter can be amended by skipping the missing values or substituting them with earlier values, while others do not need to be changed (Koopman and Ooms, 2002). Further, this property of the Kalman filter offers a solution to the forecasting challenge, since we can regard future observations as a set of missing values.

An important advantage of UC is that it does not require identifying the break points prior to estimation, since it identifies them through the recursive process. While other methods, such as the HP filter, require the researcher to arbitrarily decide on the smoothness component of the cycle λ (explained below), the UC approach lets the data decide the relative importance of the shocks to potential output and the corresponding output gap (Gerlach and Smets, 1997, Bauer et al., 2003). Kuttner (1994) finds that the UC approach gives a higher role to the transitory shocks in the real output fluctuations, relative to the shocks to the trend. Further, it permits changes over time that might arise, for example, from an increase in sample size, or methodological change in survey design.

A significant further advantage of the UC method is that it does not require other unobservable variables to estimate potential output, such as the natural rate of unemployment. What is important in the case of policymaking is that the decomposition made through the UC model is usually made based on the past and current data, so that the computation of potential output and output gap can be made in real time, and thus continuously updated (Commandeur and Koopman, 2007). On the other side, the UC method has the disadvantage of requiring considerable programming and is sensitive to the initial set of parameter values (Cerra and Saxena, 2000).

2.4.6. The Hodrick-Prescott (HP) filter

The Hodrick-Prescott (HP) filter is one of the most widely used methods for estimating potential output and the output gap. The HP filter was firstly developed by Hodrick and Prescott in 1997, as a practical alternative for decomposing macroeconomic series. Similar to previous approaches presented in this section, the HP filter also views macroeconomic time series as the sum of a growth component g_t and cyclical component c_t as presented in equation (2.1). In this framework, the growth component is assumed to be varying smoothly over time. They use several macroeconomic data, such as GNP, inflation and unemployment rate to indicate that these time series have a smoothly varying trend and a cyclical component. The main hypothesis in this study is that the growth component varies smoothly over the long-run.

The HP filter generates an estimate of the unobserved variable (potential output) using as a solution to the following minimization process:

$$\mathrm{Min}\{g_t\}^{T}{}_{t=-1}\ \{\sum_{t=1}^{T}\underbrace{(y_t - g_t)}_{\substack{\textit{Cycle}\\ \text{(Unobserved)}}}^2 + \lambda \sum_{t=2}^{T-1}[\underbrace{(g_{t+1} - g_t) - (g_t - g_{t-1})}_{\substack{\text{Change in growth rate of } \textit{Trend}\\ \text{(Unobserved)}}}]^{2}\} \tag{2.19}$$

The first part of this function denotes the difference between actual output y_t and potential output g_t, representing the output gap. The second part of the equation represents the sum of potential output variations, which is assumed to be integrated of order two I(2). The difference $g_t - g_{t-1}$ corresponds to the growth rate, since the expression here is presented with natural logarithms. Similar to equation (2.3) above, this expression also allows the trend to vary over time, however in a more smoothly (restrictive) way, whereas the cyclical component is stationary over a wide range of smoothing values (de Brouwer, 1998).

The ratio of variances between c_t and g_t is represented by λ in the HP filter function.[51] The larger the λ parameter is, the smoother the generated potential output will be and the frequencies of the variances of the cyclical component will be larger. This is because the greater the share of the variation in actual output attributed to changes in trend growth, the smaller the share attributed to the cycle.

If the cyclical component (output gap) and the second differences of the growth component (potential output) are normally distributed with independent error terms and means σ^2_e and σ^2_υ (equation (2.17)), it can be written in a state space form. Assuming that the cycle (output gap) is a white noise residual, then the HP filter written in a state space notation basically generates the same estimates as represented in section 2.4.5 (Kalman filter) .

The state space notation in HP filter represents a restricted form of the UC approach, meaning that the cyclical component (output gap) is assumed to have no structure i.e. white noise

[51] For annual data, Hodrick and Prescott suggest taking a λ equal to 100, whereas for quarterly data λ equals to 1600. These values are usually taken as given in most of the empirical studies using HP filter.

residual (as in the linear process, equation (2.1)). All the terms below indicate the same variables and parameters as indicated previously in the UC representation:

Measurement (observation) equation:

$$Y_t = z.X_t + e_t, \qquad e_t \sim N(0, \sigma^2_e) \tag{2.20}$$

$$Y_t = [1 \;\; 0] * \begin{bmatrix} g_t \\ c_t \end{bmatrix} \tag{2.21}$$

$$Y_t = g_t + e_t$$

Transition (state) equation:

$$X_t = b.X_{t-1} + v_t , \qquad v_t \sim N(0, \sigma^2_v) \tag{2.22}$$

$$\begin{bmatrix} g_t \\ \mu_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} g_{t-1} \\ \mu_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \zeta_t \end{bmatrix} \tag{2.23}$$

$g_t = g_{t-1} + \mu_t$ (a random walk with drift, where the drift term itself is a random walk) (2.24)

$$\mu_t = \mu_{t-1} + \zeta_t \tag{2.25}$$

In the state space form, the g_t stands for potential output, μ_t is the time varying drift term representing the potential output growth rate. Rewritten in this form, the HP filter can be estimated according to the maximum likelihood method by the Kalman filter. Harvey and Jaeger (1993) observed that, for the U.S. GDP, the HP filter gives a very similar trend to the one produced by fitting an unobserved components model in which the cyclical component in (2.1) is replaced by a stochastic cycle.

The HP filter is one of the most widely used methods in detrending and forecasting exercises. This is because it is quick and easy to implement, most of the software packages have it and does not require much programming compared to the UC method. Despite being used by almost every researcher in this field, the HP filter is also one of the most criticized approaches. First, though the magnitude of the λ parameters fits the macroeconomic time series in U.S., this smoothing parameter may not be appropriate for the same data of other countries, because its real value may be different. Yet, Hodrick and Prescott (1997) suggest using the same λ for all the data in the sample. However as Harvey et al., (2007) argue, different time series like GDP and investment, may exhibit different degrees of instability, thus using λ = 1600 for both of them may not yield appropriate results, thus the signal-noise ratios should vary (the ratios of the variances).[52] Another drawback of the HP filter is that it neglects structural breaks and shifts (Cogley and Nason, 1995), which is a particularly important limitation for application to the data from transition countries. Further, the HP filter may yield optimal results only when the original data are of order I(2), however this may not always the case. Indeed, this assumption may be potentially inappropriate, given that real economy variables, such as output, are typically I(1); only financial variables (e.g. stock indices etc.) are commonly I(2). Lastly, the HP filter also poses the end-sample problem, giving a higher weight to the last observations in the time series and a lower weight to the earlier observations in the time series.

Overall, univariate techniques represent quite attractive techniques for estimating potential output and the output gap, as they are quick, rather simple to implement and available in many software packages. In addition, they require only the actual output for estimations, usually the log of the real GDP, to extrapolate predictions from past characteristics of the time series. Univariate techniques are particularly convenient and appealing in the context of transition economies, given that these countries usually have a limited time span of data, or in some cases, some data are not even available (e.g. capital stock). However, all univariate techniques are somewhat ad hoc in nature and, most importantly, lack economic theory to explain the magnitude and variance of potential output and the output gap. This is because the univariate techniques rely solely on actual output and do not take into account other indicators such as the labour market conditions, which may also convey information on potential output and output gap (Gerlach and Smets, 1997). Further, a major drawback of the statistical filters is the inability to identify the turning points in the cycle as well as to distinguish between supply and demand shocks. This is because the univariate techniques assume (by default) that a trend (low frequency component) exists and so can mistakenly identify a different cycle from the one that is actually present (Anderton et al., 2014). A linear trend is already inappropriate for estimating unobserved components of time series, given that it violates the basic assumptions of the properties of the series. The HP filter requires certain assumptions to be satisfied; otherwise it leads to quasi-optimal, biased results (Harvey and Trimbur, 2008).

[52] The more attributed to the cycle, the less to the trend, but the bigger the λ, the more attributed to the trend. If that is right, then in the transition economies context a bigger λ may be preferable.

The UC approach represents a more advanced and richer structural approach, which considers some of the time series properties and does not suffer from the ‘end sample’ problem.

Multivariate methods that include structural information to extract potential output and hence, the output gap will be elaborated in the following section.

2.5. Multivariate techniques

The multivariate approaches will be explored in a more exemplified context, without the same full technical exposition, such as with equations and assumptions, as presented in the univariate techniques section. This is because, the multivariate approach requires significant technical exposition, which may not be appropriate for a PhD chapter. Examples of multivariate technical expositions can be found in Engle and Watson (1981), Mitra and Sinclair (2010), Chen and Mills, (2012), Stock and Watson (2015), etc., whereas examples of structural equations in the multivariate framework will be presented in section 2.5.2. In principle, though the objective of the multivariate methods is the same as in the univariate approaches, now unobserved components are extracted from several variables rather than one variable, conditional on other data. Multivariate techniques represent augmented forms of univariate techniques, since they exploit economic relationships to estimate potential output and the output gap. Some of the widely used multivariate techniques for estimating potential output and the output gap are the production function, the multivariate UC method, the multivariate Structural Vector Autoregression (SVAR) method, and some of the more recent ones like the Bayesian estimation of the output gap. In this section we will briefly examine the properties of the production function method and the SVAR approach, with the largest part of the discussion focussing on the multivariate UC method. The reasons for this emphasis are presented below.

2.5.1. The production function approach

The production function represents one of the oldest multivariate estimation techniques. This approach has strong theoretical foundations in economic literature (Perloff and Wachter, 1979). The production function method mostly relates to the original definition of potential output, since it exploits the relationship between inputs (labour, capital and technological change) with output, and enables to identification of the contribution of each factor to growth (De Masi, 1997). Potential output in this approach is generated when the factors of production (inputs) are fully utilized (as discussed section 2.2). However, the production function approach is also criticized on many levels. First of all, estimation relies on imposing a functional form on the production function, with the Cobb-Douglas being historically one of the most widely used. However, the Cobb-Douglas functional form holds only when a set of restrictive assumptions hold, such as constant returns to scale, which is a considerable limitation of this form. Second, technological progress is assumed to change smoothly over time, which may not be applicable in those countries undergoing intensive structural changes, such as transition countries (Stramkova et al., 2010). Moreover, the treatment of the Total Factor Productivity as a residual reduces the accuracy of the generated results. Third, this approach cannot be operationalized in the case of most transition countries, owing to the lack of data on capital stock. Even in the more advanced European transition countries where capital stock data are available, in the early stages of transition process, these data are unreliable. Fourth, the production function approach still needs to use statistical filters, such as the HP filter, to estimate the Total Factor Productivity trend, consequently bringing along with it the disadvantages of these measures (Cerra and Saxena, 2000). The production function also is not sufficiently flexible in updating real time estimates, given that the changes in productivity are rarely apparent in real time (Kuttner, 1994). Finally, growth models usually require long time series to conduct proper research, which again poses a difficulty especially in the countries with a short span of data.

2.5.2. The multivariate UC approach

The multivariate UC method represents an enhanced form of the univariate UC method, by adding economic relationships into the model. The advantage of the multivariate UC method is that this approach allows extra information to be included in the system, which via the Kalman algorithm enables more observable variables from which to extract the unobserved components. In effect, this increases the accuracy of the estimated results (potential output and output gap), relative to the case when the algorithm uses only actual output to generate the desired results. The estimation procedure is similar to the one described in section 2.4.4 and 2.4.5. First, the model is written in the state space form, where the measurement equation relates the observed variables to the unobserved variables, and in a simultaneous way, the transition equation sets a time series process for the unobserved variables. In the case of the multivariate UC, additional measurement equations will enter in the system, depending on the economic relationship we are interested to model. Second, conditional on the initial parameters (coefficients) set to define the state of the model, the maximum likelihood estimation with the Kalman filter will generate the ‘optimal estimators’, i.e. the final results.

Provided that the UC method is a flexible structure, which can handle a fair amount of structural information simultaneously, it can be used to estimate different kinds of economic relationships, such as trade balance as an external side equilibrium (Szorfi, 2011), arbitrage pricing in exchange rates (Bolland and Connor, 1997), etc. However, one of the most famous economic relationships when it comes to estimating potential output is the relationship between potential output and inflation, via the Phillips curve (Stock and Watson, 1996) or an enhanced version of Phillips Curve-Okun's law combination (Kuttner, 1994, Benes and N'Diaye, 2004).

The short-run Phillips curve implies that changes in the rate of inflation contain information about the output gap; thus this framework may be useful in extracting unobserved variables, such as potential output and the output gap (as discussed in section 2.2). One of the early applications of the Phillips curve in the framework of the signalling equations (represented in state space form), is the work of Kuttner (1994) and later Gerlach and Smets (1997). These authors augmented the Watson model discussed in section 2.4.5, by introducing the inflation variable in the (signalling) equation:

$$\Delta\pi_t = \mu_\pi + \gamma\Delta y_{t-1} + \beta c_{t-1} + v_t + \delta_1 v_{t-1} + \delta_2 v_{t-2} + \delta_3 v_{t-3} \qquad (2.26)$$

The term $\Delta\pi_t$ denotes the change in inflation, whereas Δy_{t-1} stands for the lagged GDP growth rate to capture the positive correlation between inflation and real output growth, captured by the γ coefficient. The lagged cyclical component (output gap) c_{t-1} represents an additional explanatory variable in the stochastic trend model, of the current and previous changes in inflation. The response of inflation to the output gap is captured by the β coefficient. The μ_π term is the intercept, the v_t is the white noise error term and the δ coefficients represent the moving averages of the dependent variable ($\Delta\pi_t$).

Later the Phillips curve equation was extended by adding into the equation information on the labour market and expectations formation. This led to even richer models accounting for the relationship between potential output and the NAIRU relationship, even though they became more complex and difficult to measure. For example, Appel and Jansson (1999) used five signalling equations in the multivariate UC approach, in order to extract the unobserved variables:

$$\Delta\pi_t = \rho(L)\,\pi_{t-1} + \eta(L)(u_t - u_t^n) + \omega(L)z_t + \varepsilon_t^{pc} \qquad \varepsilon_t^p \sim iid(o, \sigma^{2\,pc}) \qquad (2.27)$$

$$y_t - y_t^p = \phi(L)(u_t - u_t^n) + \varepsilon_t^{ol} \qquad \varepsilon_t^{ol} \sim iid(o, \sigma^{2\,ol}) \qquad (2.28)$$

$$u_t^n = u_{t-1}^n + \varepsilon_t^n \qquad \varepsilon_t^n \sim iid(o, \sigma^{2n}) \qquad (2.29)$$

$$y_t^p = \alpha + y_{t-1}^p + \varepsilon_t^p \qquad \varepsilon_t^n \sim iid(o, \sigma^{2\,p}) \qquad (2.30)$$

$$u_t - u_t^n = \delta(L)\,(u_{t-1} - u_{t-1}^n) + \varepsilon_t^c \qquad \varepsilon_t^n \sim iid(o, \sigma^{2c}) \qquad (2.31)$$

In this set of signalling equations, two relationships are modelled. The first equation (2.27) represents the Phillips curve relationship, which models the excess demand pressures on domestic inflation. In (2.27), the $\Delta\pi_t$ is the change in inflation, the lagged term π_{t-1} represents inertia in inflation, the terms $\rho(L)$, $\eta(L)$ and $\omega(L)$are the lag operators, the u_t is the actual rate of unemployment and the u_t^n is the NAIRU, where the $(u_t - u_t^n)$ is the unemployment gap denoting the excess demand, and the coefficient η represents the inverse relationship between the change in inflation rate with the unemployment gap. The z_t stands for the vector of supply shocks proxies. The equation (2.28) captures the inverse relationship between the output gap $y_t - y_t^p$ and the unemployment gap $(u_t - u_t^n)$. The three following equations define the evolution of the unobserved variables, where potential output y_t^p in (2.30) is represented by a random walk with drift process, whereas the NAIRU u_t^n in (2.29) follows a pure random walk. In addition, all the error terms are individually and independently distributed. In this framework, the demand side variables are proxied by the output gap, the unemployment gap, such as in (2.31), (in some other representations through the rate of capacity utilization also), whereas the supply side indictors are the potential output, the NAIRU, long-run inflation etc. This model is also solved using the Kalman filter. The estimation process is similar to the one described under the univariate technique section, however in a more complex fashion. Nevertheless, the basic principles and assumptions are the same.

Besides simultaneously estimating economic relationships in the model, the UC method has the advantage of being a flexible structure for modelling potential output and output gap, since it enables the introduction of exogenous variables to capture shocks as well as dummy variables to deal with structural breaks in the model. For example, Kichian (1999) extends the potential output UC model to account for structural breaks in the model, such as price shocks or changes in monetary regimes, by allowing a one-time change in the potential output. Other studies like Kuttner (1994) augment the Phillips curve with additional inflationary variables (core inflation, broad inflation etc) or use other country specific variables to account for shocks in the model. For example, Sramkova et al., (2010) use a more complex framework to estimate an output gap-NAIRU application for Slovakia, by using simultaneously eight signalling equations:[53]

$$y_t = g_t + c_t \qquad (2.1)$$

$$g_t = g_{t-1} + \mu_{t-1} + \varepsilon_g\,, \qquad \varepsilon_g \sim iid(o,\ \sigma^2_g) \qquad (2.3)$$

$$\mu_t = \beta\mu_{t-1} + (1 - \beta)[\gamma_1 + \gamma_2 ln(0.3 + FDICUMpc)] + \varepsilon_\mu, \qquad \varepsilon_\mu \sim iid(o,\ \sigma^2_\mu) \qquad (2.32)$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \varepsilon_c, \qquad \varepsilon_c \sim iid(o,\ \sigma^2_c) \qquad (2.10)$$

$$u_t = u_t^n + u_{t,}^c \qquad (2.33)$$

$$u_t^n = u_{t-1}^n + \varepsilon_n, \qquad \varepsilon_n \sim iid(o,\ \sigma^2_n) \qquad (2.29)$$

$$u_{t,}^c = \alpha_1\, c_t + \alpha_2\, u_{t-1}^c + \varepsilon_u, \qquad \varepsilon_u \sim iid(o,\ \sigma^2_u) \qquad (2.34)$$

$$\pi_t = \lambda_1\pi_{t-1} + \lambda_2\pi_{t+1} + \lambda_3\pi_t^{CPI} + (1 - \lambda_1 - \lambda_2 - \lambda_3)\pi_t^f + \lambda_4\, u_{t-2,}^c + \varepsilon_\pi,$$
$$\varepsilon_\pi \sim iid(o,\sigma^2_\pi) \qquad (2.35)$$

The first three equations (2.1), (2.3) and (2.10) are the same as previously presented above. The identity in (2.32) decomposes the actual unemployment rate into NAIRU u_t and unemployment gap $u_{t,}^c$. The evolution of the NAIRU in equation (2.29) is specified as a pure random walk, which is common in the literature (Apple and Jason, 1998; Benes and N'Diaye, 2004).[54] Some degree of persistence in dynamics of unemployment gap is captured by the presence of the lagged values of unemployment gap u_{t-1}^c in equation (2.34). The equation (2.32) specifies that unemployment gap u_t^c is directly linked to the output gap c_t with the elasticity parameter α_1 defined in Okun's law. Besides the backward looking inflation expectations π_{t-1}, an augmented Phillips curve with forward looking inflation expectations π_{t+1} is presented in equation (2.35), where a two-quarter lagged unemployment gap u_{t-2}^c is given as a determinant of inflation. The term π_t^f stands for imported inflation, as measured by the weighted PPI. The core CPI π_t^{CPI} enters the equation as an exogenous variable. The sum of the nominal inflation determinants, backward and forward looking expectations, imported and autonomous inflation is assumed to be equal to one, in line with the long-run vertical Phillips curve theory, i.e. the sum of all λ's equals one. The novelty of this study comes in the equation (2.32) where the FDI variable enters in the dynamics of potential output growth, since it is assumed that it captures shocks in the trend of TFP, i.e. contributes to the changes in the potential output growth rate. Estimates of $\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$ determine how the estimated time-

[53] For the full state space derivation of the model, see Sramkova et al., (2010), p. 27.
[54] Benes and N'Diaye (2004) argue that the NAIRU cannot literary follow a random walk; however it is a useful econometric assumption that NAIRU drifts over time in an unpredictable way.

varying equilibrium growth develops, given the FDI inflow impact on the technology changes in the economy. Moreover, Sramkova et al. (2010) estimate potential output and the output gap in a longer set of equations: they divide the model into two sub-models to double check their results. In sub-model one, they exclude the information on the Phillips curve, whereas in sub model two, they exclude labour market information. This may be a useful robustness check, since the estimates of potential output and output gap are sensitive to the choice of variables to include in the system, and in effect may generate different estimates (Kichian, 1999).

2.6. Summary and Conclusions

This chapter has addressed the first research question of the thesis that is elaborating the concept of potential output and the accompanying measure of the gap between actual and potential output in European transition economies, from a theoretical perspective. Initially, the definition of potential output and the output gap were elaborated in the light of mainstream theories. Further, given that these two concepts were established and implemented in models for developed countries, a discussion concerning their relevance in the transition context and importance for policymaking were also added. Given that potential output and output gap are not directly observable and given the difficulties surrounding their estimation in transition economies, a considerable part of the chapter was concerned with finding the most appropriate estimation technique. In this respect, in order to understand the nature of the unobserved variables (potential output an output gap), we presented and compared the advantages and limitations of several commonly used univariate and multivariate techniques.

In doing so, we initially elaborated the evolution of estimations of the unobserved components techniques, based on different conceptual frameworks of the trend (i.e. potential output), such are linear, deterministic and stochastic trends. Further, in order to better describe the behaviour of the observed actual output or other macroeconomic time series, other commonly used decomposing techniques were taken into account, such as the production function, the SVAR and the Hodrick-Prescott filter and unobserved components (Kalman filtering technique) approaches. After lengthy consideration, particularly concerning the characteristics of the transition economies (especially three selected countries in our thesis Kosovo Estonia and the Czech Republic), the unobserved components method operationalized via the Kalman filter was chosen as the preferred method to estimate potential output and the output gap. Since the unobserved components method represents a complex structure for estimating potential output and output gap, a detailed exposition of the univariate form was presented in theoretical and technical form. Furthermore, one contribution to knowledge of this chapter was in presenting an intuitive explanation for economists regarding the meaning, procedure and technicalities underpinning the Kalman filter method.

In addition to the univariate techniques, some of the multivariate approaches such as the production function, Structural VAR and multivariate unobserved components were also explored in a more exemplified context. While both univariate and multivariate approaches operate on the same principle, i.e. extracting the signal from the noise, the advantage of the multivariate approach stands in utilising the economic content to extract unobserved indicators, namely potential output and the output gap, thus these estimates are typically presumed to be more accurate. Nevertheless, the lack of a clear theoretical framework to be utilised in the case of transition economies when examining the output gap would still prevent us from fully employing the multivariate specification (such as the inclusion of Okun's law in the Phillips curve relationship). Therefore, in order to utilise the advantages of employing the multivariate UC approach, the Phillips curve specification will also be modified when generating our estimates of output gap, following the examples of similar studies in transition economies (Tsalinski, 2007; Bokan and Ravnik, 2012, etc.). The estimation of output gaps in three selected European transition economies employing univariate and multivariate unobserved components method will follow in the next chapter.

To sum up, each of the presented methods has its own advantages and disadvantages, and usually the choice of the method, at least for the transitional countries, is conditional on data availability. Univariate methods have the advantage of simplicity, which makes them appealing and widely used. However, they suffer from a set of drawbacks which usually makes them less accurate as compared to the multivariate methods. Multivariate methods have a distinctive advantage in that they incorporate economic relationships and potentially improve the accuracy of the estimated variables. Nevertheless, it must be stressed that no matter how sophisticated, multivariate methods still represent only theoretical models that simplify reality. Thus, over-interpretation of the obtained results is dangerous and a comprehensive analysis requires the application of more than one method and then comparison of the results.

Chapter 3. Estimating the output gap in the Czech Republic, Estonia and Kosovo

3.1. Introduction

After investigating the concepts of potential output and the corresponding output gap, their theoretical background and relevance and challenges in terms of transition context, in this chapter we move on to the estimation process. The chapter starts by briefly describing some data issues, such as data transformation and disaggregation due to the lack of appropriate frequency data. As explained in chapter 2, the unobserved components method was chosen as the preferred estimation technique in the case of European transition economies, thus estimations of potential output and the output gap will be conducted with this method. Since the technical exposition of the unobserved components approach was thoroughly presented in chapter 2, this chapter starts directly with its implementation. As was explained in chapter 1 (section 1.7) and below in section 3.2, the estimation of unobserved components will be focused on three representative European transition economies, namely the Czech Republic, Estonia and Kosovo. Since the unobserved components method represents a relatively complex method, due to several pre-tests, identification and programming procedures, it also represents a lengthy process to implement. Therefore, a more detailed description of procedures will take place along with estimating potential output and the output gap for the Czech Republic, whereas for Estonia and Kosovo the same procedure will be repeated with a shorter description.

The rest of the chapter is organised as following. Section 3.2 describes data used and their source. Section 3.3 introduces the use of univariate and multivariate approaches of unobserved components method. Section 3.4 presents unit root tests, starting parameter values of the model, interpretation of the univariate unobserved components model and the interpretation of the multivariate unobserved components model for the Czech Republic case. The same procedure will be repeated in section 3.5 for Estonia and in section 3.6 for Kosovo. Section 3.7 concludes the chapter.

3.2. Data

The estimation of potential output and the output gap will be undertaken for three country examples, namely Kosovo, the Czech Republic and Estonia. Since each country has its own economic characteristics, like their long-run potential output, fiscal stance, regulatory framework etc., it makes more economic sense to undertake a country by country business cycle analysis rather than pooling the data for all the countries together.[55] This research project is concerned with, amongst others, measuring potential output and the output gap in European transition countries, which could mean carrying out the estimations for all the 16 countries, one by one. However, estimating potential output and output gap for each of the European transition country is time consuming and brings no significant added value. So for illustrative purposes, we will carry out the estimations for three countries only (see section 1.7 for more explanation on the choice of selected countries). The data used for all the countries will be quarterly. The time span of the data is relatively short, from Q1 2002 to Q4 2013 for Kosovo and for the Czech Republic and Estonia the time series span from Q1 1998 to Q4 2013. However, most of the empirical research in the transition countries is characterised by a short time-span of data, particularly in business cycle analysis. For example, Darvas and Vadas (2003) use approximately 15 years of annual data and 12 years of quarterly data; Benes and N'Diaye (2004) use quarterly data from 2000-2009; Antonicova and Hucek (2005) use quarterly data for 1995-2004; Tsalinksi (2007) use quarterly data from 1999-2007 and Sramkova et al. (2010) use a yearly time span of 1996-2009.

Only the real GDP data is required for estimation of the univariate UC model.[56] The nominal GDP data will be adjusted with GDP deflator, to extract the real GDP data for all three countries. Apart from changes in Consumer Price Index (CPI), following Kuttner's (1994) model extensions to augment the Phillips curve equation, other explanatory variables that proxy for inflation, such as changes in food and oil prices will be used in the multivariate UC model (see also Gerlach and Smets, 1997 and Ogunc and Ece, 2004). Oil price is proxied through the average crude price. Since this series was expressed in U.S. dollars per barrel we multiplied the series with the euro per dollar exchange rate to obtain the series expressed in euro currency. The food prices are taken from international food price index FAO. The oil and food prices are the same for the three countries. All the changes will be transformed into the first differences of the log of the respective variable in order to gain the percentage change of the period. The base year of the CPI index in Kosovo is year 2002, whereas in Czech Republic and Estonia is year 2000. The inflation variables are annualised. The data for

[55] As it will be later seen in chapter 3, each country's specific characteristics will also be reflected in country specific parameters/priors.

[56] The Statistical Agency of Kosovo provides only yearly data for the GDP series, hence a disaggregated data from yearly into quarterly will be used. The disaggregation process from yearly GDP into quarterly GDP was conducted from the economists in the Central Bank of the Republic of Kosovo. The methodology used to disaggregate yearly into quarterly data was mainly based on the weighted averages of the historical data of all the sectors of economy, e.g. external and financial sectors.

Kosovo are taken from Statistical Agency of Kosovo (SAK), IMF World Economic Outlook (2015) and calculations from the Central Bank of the Republic of Kosovo. For the case of the Czech Republic and Estonia the data source will be EUROSTAT. In addition, since quarterly data are being used, all the data is this chapter are seasonally adjusted.

3.3. The univariate and multivariate unobserved component model for the Czech Republic, Estonia and Kosovo

For the purpose of identifying the main economic events (milestones, breaks etc.) and to better interpret the estimated results for potential output and the output gap, first the preliminary view of economic developments previously provided in section 1.7 for the Czech Republic, Estonia and Kosovo will be taken into account. Second, before initializing the univariate model, the GDP series will be tested for stationarity. Further, a description of the procedure for calculating and estimating the initial parameters will take place, followed by the interpretation of the univariate model results. A more detailed description of the whole process will be presented while constructing the models in the Czech Republic case. Afterwards, we will analyse in a similar way the univariate and multivariate models for Estonia and Kosovo avoiding repetition.

3.4. The output gap in the Czech Republic

3.4.1. Unit root tests of the GDP series, Czech Republic

Unit root testing is the starting point in most of the conventional time series analyses due to the fact that most time series methods have a stationarity of the data as a precondition. However, this does not refer to the UC model, as argued by Harvey and Jaeger (1993) and as discussed in Madalla (1997). Additionally, unit root tests tend to have low power in small samples, because the random walk component can have a small variance (Cochrane, 1991). Nevertheless, we decided to conduct unit root tests because they can give an indication how to structure the unobserved component model. Unit root tests are usually useful for determining the specification of the model, for example, to define the order of integration of variables (e.g. first differencing or second differencing) and the number of lags to include in the final model (Harvey and Proietti, 2005). For example, if the real GDP series is integrated of order one, this means that the trend (potential output) is also integrated of order one. Furthermore, the UC model enables modelling of the non-stationary variables alongside stationary variables and the stationarity of the series is not a prerequisite (Harvey and Jaeger, 1993, Madalla, 1997).

The real GDP series was tested for unit root, using the Augmented Dickey-Fuller (ADF) test and Phillips-Perron (PP) test. The ADF test was conducted with and without a trend, as discussed in section 2.4.2. The PP test uses Newey-West heteroskedasticity and autocorrelation robust standard errors.

H_0: The series is non-stationary (unit root):

Table 3.1. Unit root test univariate model, Czech Republic*

Country	Series	ADF Test, intercept included	ADF Test, intercept & trend included	Phillip-Perron Test	Perron test with structural breaks	Decision
Czech Republic	Log-levels of quarterly real GDP	Not rejected at any level of significance.	Not rejected at any level of significance.	Not rejected at any level of significance.	Not rejected at any level of ignificance.	Non-stationary
	Δ Log GDP	Not rejected at 1%, rejected at 5% level of significance.	Rejected at 10% level of significance.	Rejected at all levels of significance.	-	Possibly Stationary

*Note: If the trend in ADF (intercept+trend) test turned out as insignificant, the ADF (intercept only) and Phillip-Perron test were considered to judge a series as stationary or not.

An important property of our model is the presence of the unit root in the real GDP data for all three countries, which is confirmed for the Czech Republic through the ADF tests for non-stationarity, as reported in table 3.1 (see Appendix 3.1 for the full output of results).[57] This implies that the real GDP, as are many other macroeconomic variables, is integrated of order one, I(1). The presence of a unit root ($\rho = 1$) also implies that the trend component should be specified as a non-stationary stochastic component and that we should use difference GDP rather than level GDP. As discussed in section 2.4.3, this may be a more appropriate trend specification, especially for transition countries, as it allows capturing movements in the trend of real GDP. If the trend (potential output) is I(1), this means that the drift term (see section 2.4.4 for Watson specification) should be specified as constant through time (implying a smoothly varying trend over time) and the trend (potential output) as a random walk with drift. Otherwise, a time varying drift also means that the series should be modelled as integrated of order two, I(2), (Harvey and Jaeger, 1993). Even if we assume a slowly evolving growth rate for the Czech Republic case, this assumption may not be plausible in a dataset of only 17 years. In addition, we will add a dummy variable to capture the structural break in the fourth quarter of 2008 caused by the Global Financial Crisis (GFC), hence the equation of potential output will take the following form:

$$g_t = \mu_g + g_{t-1} + \lambda * d2008 + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma^2_\varepsilon) \qquad (3.1)$$

[57] For more details on Perron unit root tests with structural breaks see section 5.6.

where the dummy models a one period break in Q4 2008, and λ is its coefficient.[58]

3.4.2. Starting parameter values of the univariate UC model, Czech Republic

In order to proceed with the estimations using the Kalman filter, we should first decide on the choice of the initial parameter values (μ_g, ϕ_1, ϕ_2, σ^2_ε and σ^2_υ) and the starting values of the trend and cycle (y_t and c_t). Unfortunately, the initial parameter values are usually unknown and need to be derived (estimated, assumed or calibrated) from outside the model and then fed into the Kalman filter to estimate the model. The number of the initial parameter values in an UC model using Kalman filter is usually larger compared to other models (Table 3.2).

Table 3.2. The parameter estimates using different approaches

Linear Time Trend Regression	Hodrick-Prescott Filter	Unobserved Components
Intercept α *Slope* β *Variance* σ^2	*Smoothness parameter* λ *Variance* σ^2_e *Variance* σ^2_v	*Starting trend* y_t *value* *Starting* c_t *value* *Drift terms* μ_g *Autoregressive coefficient* ϕ_1 *Autoregressive coefficient* ϕ_2 *Trend Variance* σ^2_ε *Cycle Variance* σ^2_v

There are two semi-formal ways of deriving initial parameter values. First, one can use a HP filter to decompose the actual GDP into trend and cycle and use them as starting values. Afterwards, one can regress trend and cycle as dependent variables in an OLS regression and use those variances as starting values (Hamilton, 1994, de Brouwer, 1998). Another way is to assume that the starting value for potential output is equal to the starting vale of actual output and assume that the output gap and other parameters are zero, or simply make arbitrary guesses about all the initial values (Hamilton, 1994). Alternatively, if no prior knowledge is available for the initial conditions, EViews software estimates the starting values by default. It

[58] In the ADF-Perron test with structural breaks, the impulse dummy was significant at 1 percent level of significance, whereas the intercept-shift dummy was significant only at 10 percent levels of significance. However, the estimated intercept-shift dummy turned out as insignificant so it was dropped out of the equation.

estimates OLS regressions for the measurement equation using the first few observations and holding the initial mean vector fixed. Such obtained parameters are further treated as initial values for the parameters. However, this method rarely leads to model convergence, especially if the series is short, therefore researchers should undertake a lengthy process of trial and error with initial parameters until the model eventually generates ‘reasonable’ results. The initial parameter values should be such that they will minimize the difference of the mean squared error on one side and maximise the log-likelihood function as much as possible on the other side (see equation 2.17 in chapter 2).[59] However, the maximum likelihood function in the case of short time-series datasets, such is the case with many transition economies, usually poses problems, amongst others, in convergence of the model (even though convergence is not the only criteria to use to assess model validity, however if the model does not converge it cannot be estimated). Thus, using totally unrestricted parameter values does not seem reasonable. Hence, when using small datasets, particularly those of the transition economies, Darvas and Vadas (2003) suggest using a combination of the following criteria as guidelines for judging the results to be ‘reasonable’:

a) Convergence of the model (after a set of iterations),
b) Statistical validity (serially independent residuals, normally distributed residuals and homoscedastic variances),
c) Historical events (shocks, shifts in regimes, etc.), and
d) Reasonable results/Judgemental element (e.g. the estimated output gap should not be higher than 15 percent).[60]

If the first combination of initial values does not lead to a ‘reasonable’ model, then, as suggested by Hamilton (1994) we should proceed with alternative (trial and error) guesses in order to maximise equation 2.17.

In the following, we will elaborate the set of each initial parameter value in our model in more detail. First, in the random walk models, the best predictor of the drift term μ_g is usually the average growth rate of actual output. Second, since we are modelling a non-stationary model,

[59] Darvas and Vadas (2003, p.41) explain that the Kalman filter allows for the maximum likelihood function to estimate the parameters, whether the actual output y_t and the error ε_t term are stationary or not.

[60] Darvas and Vadas, (2003, p. 11) argue that “At the first glance, this seems to be too high but the early years of transition requires such a high value. On the other hand, several specifications failed to fulfil this criterion, for example, some provide more than 30% output gaps”. This may also be the case with transition economies, which did not utilize their potential for a fairly long period of time.

we cannot rely on unconditional mean and variances (these apply to stationary models), so both variances are allowed to vary.[61] Clark (1987) suggests that the potential output standard error to output gap standard error ratio equals around 0.9. Using Clark's ratio, we have weighted our default potential output standard error by 0.9, and left the output gap standard error unrestricted. Third, the autoregressive coefficients (ϕ_1 and ϕ_2), should be set to fulfil the stationarity criteria, that is the sum of the two coefficients should not equal one: $\phi_1 + \phi_2 \neq 1$ (Harvey, 2005). Compliance with this rule is necessary for two related reasons: first, if the sum of the autoregressive coefficients equals one, then the model does not converge; second, the cycle may contain a unit root. Kuttner (1994) suggests using $\phi_1 = 1.4$ and $\phi_2 = -0.5$. Darvas and Vadas (2003) on the other hand suggest to take autoregressive coefficients in the range of - 0.3 or + 0.3 of the default estimates. Last, the mean, variance and the covariance of the model should be set. Following the literature, the mean and covariance should be set to zero and the variance equal to a very large number, reflecting the uncertainty of the starting value (Koopman et al., 1998; Morley et al., 2002). Given that a State Space model by definition assumes independent error terms, EViews by default assumes that the covariance between error terms is zero (Quantitative Micro Software, 2010).

We have set the starting values (μ_g, ϕ_1, ϕ_2, σ^2_ε, σ^2_υ and λ) for the univariate UC model for the case of the Czech Republic in the following manner. The initial value of the potential output was taken from the first period of the HP filtered trend. The autoregressive coefficients ϕ_1 and ϕ_2 were set in the usual manner as suggested by theory (Clark, 1989 and Kuttner, 1994). The variance of the potential output was generated by running an OLS regression into a HP-filtered potential output. The variance of the output gap was generated in a similar fashion. In addition, the initial dummy variable coefficient λ were set by running an OLS regression from equation 3.3 in section 3.4.3.

The initial parameter values for the Czech Republic univariate model are shown in Table 3.3:

[61] If both variances are zero ($\sigma^2_\varepsilon = 0$ and $\sigma^2_\upsilon = 0$) then the trend is deterministic; however, allowing the σ^2_υ to be positive and σ^2_ε to be zero, the estimated trend will be relatively smooth (Harvey, 2005).

Table 3.3. Initial parameter values of the univariate UC model, Czech Republic

μ_g	σ^2_ε	σ^2_v	ϕ_1	ϕ_2	λ
0.01	0.928	0.5	1.4	-0.5	-0.08

3.4.3. Interpretation of the univariate UC model, Czech Republic

The UC model should satisfy three properties of the standardised prediction errors (one step ahead errors divided by their variances), these are i) *independence*, ii) *homoscedasticity* and iii) *normality*. Serially independent residuals are the most important assumption in a UC model, which was tested using the Box-Ljung test with 10 lags (Table 3.4). This test suggests that the null of no serial correlation between the residuals of the output gap cannot be rejected. The normality of the residuals was tested using the Jarque-Bera test, which suggests a normal distribution of the residuals. Last, the variance homoscedasticity of the residuals is not available in EViews for state space models.

As can be seen in table 3.4, the univariate model for the Czech Republic satisfies the assumption of serially independent residuals. However, the residuals are not normally distributed, based on the Jarque-Berra test for normality, probably due to some more extreme value in Q4 2008 and Q2 2010.[62] The estimated coefficients of the model, via Kalman filter and the maximum likelihood function are as presented in table 3.4 (Appendix 3.2).

Table 3.4. The estimated parameter values of the UC univariate model, Czech Republic

Parameter estimates						Diagnostics	
μ_g	σ^2_ε	σ^2_v	ϕ_1	ϕ_2	λ	Serial correlation Q(10)	Normality
0.34***	0.109***	0.000	1.143***	-0.109**	-1.758	0.443	0.000

Note: The log-likelihood function of the model is 24.97. The model converged after 9 iterations.

***Significant at 1% level of significance.

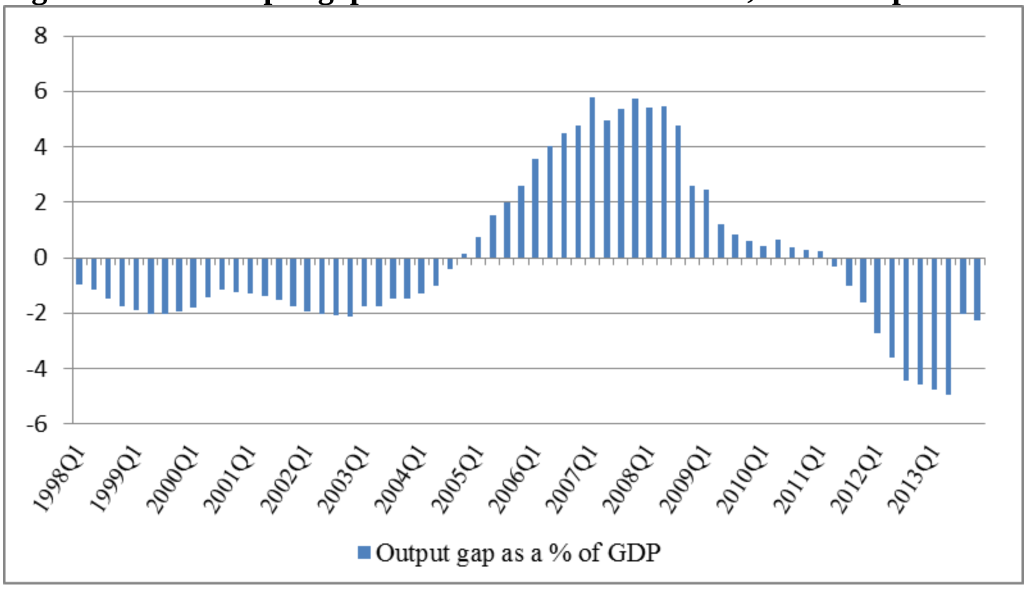
** Significant at 5% level of significance.

The estimated potential output growth rate μ_g is around 0.3 percent per quarter, or 1.2 percent annually, which is lower than the actual output growth rate (see figure 3.3), meaning that the estimated result suggests a slower potential output growth. The dummy variable for the year Q4 2008 turned out to be statistically insignificant, suggesting that the shock from the crisis

[62] It may be useful to recall that from the three properties that the residuals of an UC model should satisfy: independent serially correlated residuals, homoscedastic variances and normal distribution of the residuals, normal distribution of the residuals is the least important criteria for the model to be considered as valid. However, non-normality does not cause in-efficient estimators.

had only a transitory impact on the potential output.[63] The variance of potential output $\sigma^2\varepsilon$ turned out as highly significant, suggesting that most of the variations in the real GDP of the Czech Republic are secular than cyclical. Besides, the magnitude of the potential output variance suggests that the model has probably identified more than one structural break in the series (probably austerity measures introduced in late 1990s or GFC in 2008). On the other side, the variance of the output gap σ^2_v does not statistically differ from zero and turned out as insignificant, again confirming that secular movements are more important in actual output. Both autoregressive coefficients ϕ_1 and $\phi_{2,}$ turned out as highly significant indicating the output gap in the Czech Republic follows an AR(2) process.

Figure 3.1. The output gap of the univariate UC model, Czech Republic[64]



The estimated output gap presented in figure 3.1 varies most of the time between -4% to +4% of GDP.[65] The highest value of the output gap (+5.79 percent of GDP) was depicted in Q1 2007, whereas the lowest value of negative output gap of -4.58 percent of GDP was in the last quarter of 2012. As suggested from figure 3.1., at the very beginning of our sample (1998) the output gap of the Czech Republic was negative, coinciding with a period after the currency crisis, and followed by a period of negative growth rate of the actual output, until the

[63] Even after the dummy variable *d2008* turned out insignificant, it is still informative regarding its impact on the potential output. Moreover, adding the dummy variable helped the model to converge. As we will see later, the dummy variable for 2008 is statistically significant in the multivariate model.

[64] Since our research is primarily concerned with the output gap rather than potential output in European transition economies, potential output is not presented in respective figures. Furthermore, as we will see in chapter 5, for simplicity the potential output is assumed as constant. to fluctuate in rather longer-term,

[65] Positive values of the output gap indicate that the economy is operating above the potential output, whereas negative values indicate that the economy is operating below its potential.

beginning of the 2000s. Based on the estimated output gap, the expansion of the economy returned in the last quarter of 2004, which represents a period of high inflows of foreign investments and an export boom. The late impact of the global crisis is reflected in cycle of the Czech economy arriving only in the third quarter of 2009, putting the positive output gap at a fairly low base (around 0.5 percent of GDP) until the recession starts in 2011 onwards.

Even though the univariate model for Czech Republic has provided us with a fairly clear picture of the business cycle during the last 15 years, we will continue estimating the potential output and the output gap with economic content in the next section.

3.4.4. Multivariate UC model, Czech Republic

As discussed in previous sections, multivariate models provide a richer structure and utilize additional content form economic theory, for the purpose of increasing the accuracy of the unobserved variables, potential output and the output gap. We will try to increase the accuracy of the estimated unobserved variables, by adding an additional measurement (signal) equation into to the previously presented model that is a Phillips curve equation. As in the univariate specification, a dummy variable denoting the last quarter of 2008 will be added into the potential output equation. Further, it is worth noting that the Czech Republic conducts an inflation targeting regime, hence the Phillips curve will be modified by adding inflation expectations to the inflation equation. This way, the Phillips curve will depend on inertia (π_{t-1}), inflation target proxied by inflation expectations, and the output gap:

$$y_t = g_t + c_t \qquad (2.1)$$

$$\pi_t = a(L)\,\pi_{t-1} + (1-a)*\pi_{targeting} + b(L)D_t + c(L)z_t + e_t \quad e_t \sim iid(0, \sigma^2) \qquad (3.1)$$ [66]

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \upsilon_t \qquad \upsilon_t \sim iid(0, \sigma^2_\upsilon) \qquad (2.10)$$

$$g_t = \mu_g + g_{t-1} + \lambda * d2008 + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma^2_\varepsilon) \qquad (3.3)$$

As discussed in the theoretical part of chapter 2, Phillips curve represents an essential structural relation that may convey information about the potential output and the output gap in the economy. Shepherd and Driver (2003) argue that ”the Phillips curve can be expressed as a relationship between price inflation and cost inflation, or after some manipulation, a

[66] It should be noted that since Czech Republic conducts inflation targeting, a HP-filtered inflation trend was extracted to account for inflation expectations was also included in the Phillips curve equation - $(1- a_1)\ \pi_{targeting}$. However, whenever this variable was included the model did not converge so we had to drop it from estimation.

relationship between inflation and its own lagged values". As argued by Gordon (1997), in an amended Phillips curve equation we can capture supply and demand-side shocks that may impact the unobserved variables. The modified Phillips curve, equation 3.1, includes the demand side pressures D_t, usually captured by the output gap, the unemployment and the capacity utilization rates. The π_t term stands for inflation, π_{t-1} stands for the lagged effect during the adjustment process of prices (inertia); L is for the lag polynomial, since each variable can be modelled with a different lag structure. The a and b terms are the parameters of the lagged inflation and demand side variables. As argued in orthodox theory, changes in demand variables are those that affect the inflation rate, thus they need to enter into the equation as first differences or as the first lag of the levels. Further, we include exogenous supply-side shocks variables z_t which are usually proxied through the change in import prices, change in food prices, change in oil prices, change in the real exchange rate etc. (Gordon, 1997), c is the parameter on the supply-side variables (in our case the output gap) and λ is the parameter of the dummy variable that controls for the GFC effect.

In order to formalise the Phillips curve equation, we firstly need to decide on the variables to include in the equation. We choose the output gap as a proxy for demand side pressures (Basarac et al., 2011), and the changes in food and oil prices as supply side shocks. The inflation equation for the Czech Republic is as follows:

$$\pi_t = \mu_\pi + a(L)\pi_{t-1} + b_1 c_t + b_2 c_{t-1} + d_1 \Delta FP_t + d_2 \Delta OP_t + d_3 \Delta GDP_{t-1} + e_t \quad e_t \sim iid(o, \sigma^2) \qquad (3.4)$$

where, μ_π is the constant term, the a term is the coefficient of the lagged inflation, b_1 and b_2 are the coefficients of the output gap and lagged output gap, respectively. The ΔFP, ΔOP and ΔGDP_{t-1} stand for the changes in food prices, changes in oil prices and growth rate of the GDP in previous quarter, with their respective coefficients d_1, d_2 and d_3. The GDP growth rate is entered into the equation in order to capture the positive relationship between inflation and GDP growth. The e_t is the white noise error term with mean zero and variance σ^2.

The inertia variable is constructed as an ARIMA process, following the Box-Jenkins approach. The Box-Jenkins approach involves a three step procedure: identification, estimation and diagnostic checking and forecasting (Box and Jenkins, 1970). We will deal with the first two stages of choosing the best fitting ARIMA model, whereas the forecasting stage is outside the scope of this research project. The identification stage involves testing the series for stationarity and a tentative autoregressive moving average order identification. In the estimation stage, we estimate the previously identified series (level or first differencing series) and decide on the order of the ARMA process for the inflation variable. Further, this stage also involves diagnostics checks which help in a better judgement on deciding for the adequacy of the model. However, it should be noted that one or more ARIMA models may be appropriate.

Following the Box-Jenkins approach, in the identification step each inflationary variable will be tested for stationarity (Table 3.5).

Table 3.5. Unit root tests for inflation variables, Czech Republic*

Country	Series	ADF Test, intercept	ADF Test, intercept & trend	Phillip-Perron Test	Decision
Czech Republic	**Log CPI**	Rejected at all levels of significance.	Not rejected at 1%, rejected 5% and 10% levels of significance.	Not rejected at 1%, rejected 5% and 10% levels of significance.	Possibly Non-Stationary
	Δ Log CPI	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	Stationary
Exogenous Phillips-Curve Variables (all countries)	**Log Food Prices**	Not rejected at any level of significance.	Not rejected at any level of significance.	Not rejected at any level of significance.	Non-stationary
	Δ Log Food Prices	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	Stationary
	Log Oil Prices	Not rejected at 1% and 5%, rejected 10% level of significance.	Not rejected at 1%, rejected at 5% and 10% levels of significance.	Not rejected at all levels of significance.	Possibly Non-Stationary
	Δ Log Oil Prices	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	Stationary

*Note: If the trend in ADF (intercept+trend) test turned out as insignificant, the ADF (intercept only) and Phillip-Perron test were considered to judge a series as stationary or not.

As indicated in table 3.5 (Appendix 3.1.1 and 3.1.2), there is evidence of a unit root presence in all three inflation variables, whereas their first difference turns stationary. Hence, the inflation variables will be treated as integrated of order one, I(1).

In the second step, the choice is based on the examination of the autocorrelation (ACF) and the partial autocorrelation function of residuals of the inflation series (PACF), which are commonly used to identify the ARMA model. In the case of the geometric progression decline of the ACF and an abrupt decline of the PACF after the lag *p*, an AR(p) model should be considered. Conversely, if the ACF declines abruptly after the lag *p*, whereas the PACF declines with geometric progression after the lag *p*, than an MA(p) model should be considered. However, if both, ACF and PACF tail off, then using an ARMA model is suggested. So, firstly we estimated the In trying to choose between an AR(p) vs. MA(q) inflation model for the Czech Republic case, the correlogram of the equation (3.4) did not provide clear evidence, since both, the ACF and PACF fell abruptly after lag one. Thus, we have estimated an AR(1) and MA(1) model, meaning that we include one autoregressive term ($L=1$) of the lagged inflation (the term $a(L)\pi_{t-1}$ in equation 3.4), and check the diagnostics tests. An ARIMA model may be considered as valid when the model is free from residual correlation and has the smallest information criteria (Verbeek, 2000). The Ljung-Box test is commonly used in testing ARIMA modelling for residual correlation using Q(10)-test. Additionally, the smallest Akaike and Schwarz criteria aid in choosing the best fit model. It is worth noting that when dealing with small samples, the Schwarz criterion may be more powerful (Verbeek, 2000). We have firstly tested a pure autoregressive model of order one, by adding only one autoregressive Log-CPI variables into the equation, then following Kuttner (1994) we added one by one other independent variables (Kuttner, 1994), (table 3.6):

Table 3.6. Estimated AR(1) inflation model, Czech Republic

	Parameter estimates						Diagnostics		
	μ_π	a_1	a_2	d_1	d_2	d_3	*Akaike i.c.*	*Schwarz i.c.*	*Q(10) p-values*
1	-0.344	-0.460					3.770	3.905	0.192
	(0.787)	(0.151)							
2	-0.045	0.108***	0.069***	0.031**			0.970	1.139	0.212
	(0.070)	(0.036)	(0.025)	(0.016)					
3	-0.358	0.408**	0.358	0.092	0.008		3.477	3.698	0.064
	(0.555)	(0.167)	(0.226)	(0.070)	(0.033)				
4	-0.542	0.377**	0.386*	0.090	0.001	0.479	3.473	3.731	0.151
	(0.569)	(0.171)	(0.227)	(0.070)	(0.008)	(0.345)			

Note: Standard errors in parentheses.
*** Significant at 1% level of significance.
** Significant at 5% level of significance.
* Significant at 10% level of significance.

In table 3.6, a pure AR(1) model was estimated with one lag of the CPI inflation, together with the inflation expectations denoted by the coefficient a_2, then other explanatory variables were added one by one (Appendix 3.3). All four equations are well specified regarding the serial correlation of the residuals. However, the information criteria suggest that equation two is the best fit model for the inflation equation. In this equation, the drift term, the third lag of the CPI, changes in food prices and changes in oil prices resulted as insignificant, so that will allow us to drop them. However, inflation expectations, as expected were statistically significant.

Turning to the MA(1) model in table 3.7, the information criteria suggest that equation 3 is the best model for the inflation equation, where the drift term, inertia, inflation expectations and changes in food prices are statistically significant. The change in oil prices variable was insignificant, so we can drop it from the equation (table 3.7).

Table 3.7. Estimated MA(1) inflation model, Czech Republic

	Parameter estimates						Diagnostics		
	μ_π	a_1	a_2	d_1	d_2	d_3	*Akaike i.c.*	*Schwarz i.c.*	*Q(10) p-values*
1	-0.198	-0.575	0.699*				3.797	3.930	0.027
	-0.609	(0.759)	(0.373)						
2	-0.184**	0.723**	0.149***	0.189***			3.498	3.665	0.229
	(0.081)	(0.075)	(0.045)	(0.045)					
3	-0.410**	0.671**	0.261***	0.147***	0.016		3.165	3.384	0.349
	(0.196)	(0.085)	(0.101)	(0.044)	(0.019)				
4	-0.520**	0.672**	0.292***	0.138***	0.004	0.015	3.184	3.439	0.309
	(0.193)	(0.103)	(0.103)	(0.048)	(0.005)	(0.096)			

Note: Standard errors in parentheses.
*** Significant at 1% level of significance.
** Significant at 5% level of significance.
* Significant at 10% level of significance.

After considering the AR(1) and MA(1) models, the information criteria are smaller for the AR(1) model, hence the inflation equation will take the following form:

$$\pi_t = a_1\pi_{t-1} + b_1c_t + b_2c_{t-1} + d_1\Delta FP + e_t \qquad e_t \sim iid(o, \sigma^2) \tag{3.5}$$

where a_1 denotes the inflation inertia coefficient, the c_t and c_{t-1} represent the output gap and the lagged output gap generated from the HP filter, whereas ΔFP represent changes in food prices. In addition, the drift term of the inflation was insignificant in equation two AR(1) so we can drop it from the Phillips curve equation.

The starting values of the autoregressive coefficients and the variances of potential output and the output gap remain the same as in the univariate model. The starting values for the inflation equation were obtained using a combination of estimations as in the univariate model, judgement, trial and error and the default procedure.

The initial parameter values and the estimated ones from the multivariate UC model are presented in table 3.8 (Appendix 3.4).

Table 3.8. The starting and estimated multivariate UC model coefficients, Czech Republic

Starting parameter values		Estimated coefficients
μ_g	1.0	0.310***
		(0.065)
σ^2_ε	0.928	0.021***
		(0.385)
σ^2_v	0.5	0.024***
		(0.592)
ϕ_1	1.4	1.855***
		(0.092)
ϕ_2	-0.5	-0.876***
		(0.096)
λ	-0.08	-1.127**
		(0.570)
a_1	0.952	1.074***
		(0.157)
b_1	0.169	-.107
		(1.519)
b_2	-0.201	-0424
		(1.608)
d_1	0.118	-0.034
		(0.249)
σ^2_π	0.222	1.382***
		(0.146)
Output equation		
H0: No serial correltation:		Q(10) = 0.114
H0: Normality in the residuals:		Jarque-Berra = 0.000
Inflation equation		
H0: No serial correltation:		Q(10) = 0.130
H0: Normality in the residuals:		Jarque-Berra = 0.000

Note: Standard errors in parentheses The log-likelihood function of the model is 137.78.
The model converged after 254 iterations.

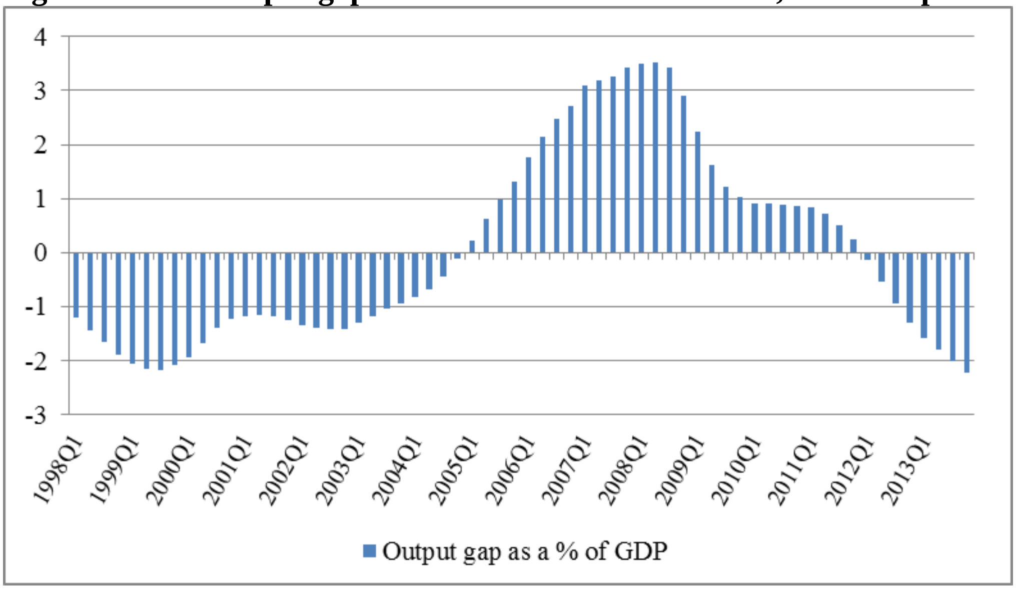
*** Significant at 1% level of significance.
** Significant at 5% level of significance.
* Significant at 10% level of significance.

The model is well specified with regard to the serial correlation, however as in the univariate model, the residuals are not normally distributed. In the case of the Czech Republic, the multivariate coefficients are more informative than the ones in the univariate model.

The 2008 dummy variable was statistically significant at 5 percent level of significance, suggesting that the GFC shock had a significant but rather temporary effect on potential output during the following time period modelled.[67] The effect of the crisis in potential output in Czech Republic case was reflected via lower inflows of foreign direct investment.

One of the most important relations in the UC multivariate model is the relation of the output gap and inflation. The relation between output gap and the inflation is negative in the case of the Czech Republic, which is in line with New Keynesian Phillips curve theory.[68] A 1 pp increase in the output gap, on average, would result in 0.2 pp decrease in the inflation rate, ceteris paribus, suggesting that a negative output gap is an indication of lower inflationary pressures. However the output gap coefficient was insignificant. Inflation inertia is highly significant with a relatively large magnitude, confirming that the previous period inflation rate tends to also persist in the current period.

Figure 3.2. The output gap of the multivariate UC model, Czech Republic



As in the case of the univariate model, the beginning of the sample period starts with a small positive output gap. It is worth noting that this crisis did not have a permanent shock on future potential output (Figure 3.7). The beginning of a positive cycle does not start until Q1 2005

67 The IMF (2012a, p. 2) suggests that "the crisis may impact the potential output in several ways. Decline in investment—due to lower profits, worsened business prospects, rising interest rates in the case of large fiscal deficits, credit tightening, slows down capital accumulation, while bankruptcies and restructuring negatively affect productivity of the already existing stock. Meanwhile, increases in unemployment may lead to higher natural rate of unemployment as longer unemployment spells lead to discouragement and deterioration of already existing skills. Lower profits, leading to lower R&D spending, would negatively affect future productivity".

68 As discussed in chapter 2, the New Keynesian theory implies that the output gap, the deviation of the actual output from its natural level due to nominal rigidities, drives the dynamics of inflation relative to expected inflation and lagged inflation (Zhang and Murasawa, 2011).

(in the univariate model it starts in Q4 2004) and the positive output gap is larger, reaching up to 9.8 percent of the GDP, whereas in the univariate model the positive output gap was not higher than 5.8 percent during 2005-2011. The negative output gap during late 1990s and the positive output gap largely coincide with IMF's estimations for Czech Republic (IMF, 2012a).The significant structural break dummy controlling for the impact of GFC, is important, since it suggests that the potential output suffers a supply-side shock in the last quarter of 2008. A second important point to emphasise is the business cycle turning point in the same period, which is in not in line with the main assumptions of the model that suggest that shocks affecting potential output and the output gap should not be correlated. After 2009, the negative output gap is consistent with the slow and even negative growth rates during this period.

In the following section, the estimations for potential output and output gap in Estonia will be presented.

3.5. The output gap in Estonia

3.5.1. Unit root test for GDP series, Estonia

As noticed in the table 3.9, the log-levels GDP variable for Estonia has a unit root, whereas its first difference is stationary, so we treat Estonian GDP as integrated of order one variable- I(1).

H_0: The series is non-stationary (unit root):

Table 3.9. Unit root tests of the univariate model, Estonia*

Country	Series	ADF Test, intercept included	ADF Test, intercept & trend included	Phillip-Perron Test	Perron test with structural breaks	Decision
Estonia	**Log-levels of quarterly real GDP**	Not rejected at any level of ignificance.	Not rejected at any level of ignificance.	Not rejected at any level of ignificance.	Not rejected at any level of ignificance.	Non-stationary
	Δ Log GDP	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	-	Stationary

*Note: If the trend in ADF (intercept+trend) test turned out as insignificant, the ADF (intercept only) and Phillip-Perron test were considered to judge a series as stationary or not.

This means that, the potential output will be modelled as random walk with a constant drift. Considering that Estonia experienced the effects of the global financial crisis at the end of 2008, a dummy variable will also is added into the potential output equation:

$$y_t = g_t + c_t \tag{2.1}$$

$$g_t = \mu_g + g_{t-1} + d2008 + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma^2_\varepsilon) \tag{3.3}$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \upsilon_t \qquad \upsilon_t \sim iid(0, \sigma^2_\upsilon) \tag{2.10}$$

3.5.2. Starting parameter values of the univariate UC model, Estonia

The univariate model for Estonia set up the same way as in the Czech Republic case, was estimated using the following initial parameter estimates (μ_g, ϕ_1, ϕ_2, σ^2_ε, σ^2_υ and λ) in table 3.10:

Table 3.10. Starting parameter values of the univariate UC model, Estonia

μ_g	σ^2_ε	σ^2_v	ϕ_1	ϕ_2	λ
0.016	0.226	0.245	1.4	-0.5	-0.326

3.5.3. Interpretation of the univariate UC model, Estonia

As can be seen in table 3.11, the model is not misspecified with regard to serial correlation, however normality test fails most probably due to some outliers in Q4 1998 and Q4 2008.

Table 3.11. Estimated parameter values of the univariate UC model, Estonia

μ_g	σ^2_ε	σ^2_v	ϕ_1	ϕ_2	λ	Serial correlation Q(10)	Normality
0.561***	0.354	0.283	1.769***	-0.792***	0.094	0.813	0.000

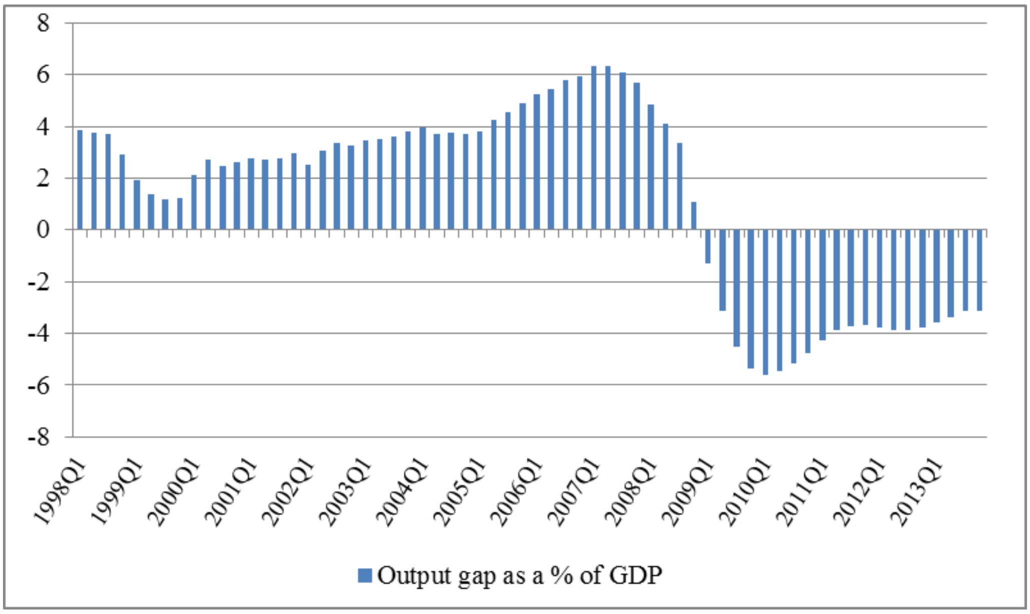
Note: The log-likelihood function of the model is 81.55. The model converged after 61 iterations.

***Significant at 1% level of significance.

** Significant at 5% level of significance.

The estimated potential output growth rate μ_g is 0.56 percent, quarterly, which is significantly lower than the actual output growth rate. The drift term was highly significant confirming that the potential output in Estonia follows a random walk with a drift process. Also, both autoregressive coefficients (ϕ_1, ϕ_2,) were highly significant confirming that the output gap follows an autoregressive process of order two, AR(2). Regarding the variances (σ^2_ε and σ^2_υ), neither were significant, meaning that at this stage it is not clear which component to attribute the variations in the actual output. Somewhat strange, considering the crisis in that period is also the insignificant dummy variable of 2008 controlling for the GFC impact.

Figure 3.3. Potential output and output gap of the univariate UC model, Estonia



Turning to the estimated unobserved components, the largest output gap stands at +6.3 percent of GDP in Q1 2007, which coincides with the highest peak of the overheating period, and the lowest output gap of -5.3 percent of GDP in Q1 2010 (Figure 3.3). It appears that the output gap was rather small at the beginning of the period, around 0.3 percent of GDP. The robust and fast catching-up period (from Q4 1998 to Q3 2008) resulted in a prolonged overheating period, which slows down only in the last quarter of 2008. The beginning of 2009 marks a turning point in the business cycle, when the deep recession starts due to the crisis.
Regarding potential output, it follows a fairly smooth path. The crisis at the end of 2008 appears to have affected potential output in Estonia only temporarily, as the trend path soon gets back to the previous one. Since the crisis resulted in a negative output gap, whereas its impact on potential output was only transitory, it may be concluded that shocks affecting potential output and output gap are uncorrelated for the case of Estonia, which is in line with the main assumption of the UC model.

In search of a better model for Estonia's unobserved components, we estimate the multivariate model in the next section.

3.5.4. Multivariate UC model, Estonia

The construction of the multivariate UC model was thoroughly explained in section 3.4.3, thus in the case of Estonia we will directly present the Phillips curve equation and briefly assess which ARIMA model fits the inflation equation better. The rest of the model specification will be similar to the Czech Republic case.

Table 3.12. Unit root tests for inflation variables, Estonia*

Country	Series	ADF Test, intercept	ADF Test, intercept & trend	Phillip-Perron Test	Decision
Estonia	Log CPI	Not rejected at any level of significance.	Not rejected at 1% and 5%, but rejected at 10% level of significance.	Not rejected at any level of significance.	Non-stationary
	Δ Log CPI	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	Stationary
Exogenous Phillips-Curve Variables (all countries)	Log Food Prices	Not rejected at any level of significance.	Not rejected at any level of significance.	Not rejected at any level of significance.	Non-stationary
	Δ Log Food Prices	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	Stationary
	Log Oil Prices	Not rejected at 1% and 5%, rejected 10% level of significance.	Not rejected at 1%, rejected at 5% and 10% levels of significance.	Not rejected at all levels of significance.	Possibly Non-Stationary
	Δ Log Oil Prices	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	Stationary

*Note: If the trend in ADF (intercept+trend) test turned out as insignificant, the ADF (intercept only) and Phillip-Perron test were considered to judge a series as stationary or not.

The results of table 3.12, suggest that the inflation variables are integrated of order one, I(1). Next, the residual correlogram on inflation series suggested that either an AR(3) model or MA(3) model may be the best fit:

$$\pi_t = \mu_\pi + a(3)\pi_{t-1} + b_1 c_t + b_2 c_{t-1} + d_1 \Delta FP_t + d_2 \Delta OP_t + d_3 \Delta GDP_{(t-1)} + e_t \quad e_t \sim iid(o, \sigma^2) \qquad (3.4)$$

The information provided in table 3.13 below suggests that the third equation is the best model fit for the inflation equation. In this equation the third lag of inertia (π_{t-3}) is insignificant so we may continue with the first two lags only. Also, the exogenous variables such as the changes in food prices and changes in oil prices were statistically insignificant so we may also drop them.

Table 3.13. Estimated AR(3) inflation model, Estonia

	Parameter estimates							Diagnostics		
	μ_π	a_1	a_2	a_3	d_1	d_2	d_3	*Akaike i.c.*	*Schwarz i.c.*	*Q(10) p-values*
1	0.012***	0.388***	0.231*	-0.311*				-5.574	-5.401	0.338
	(0.004)	(0.129)	(0.131)	(0.178)						
2	0.011	0.366***	0.270**	-0.276*	0.0003***			-5.616	-5.408	0.343
	(0.004)	(0.122)	(0.129)	(0.132)	(0.0002)					
3	0.009*	0.328**	0.332**	-0.225	0.0002	0.0001		-5.538	-5.276	0.855
	(0.005)	(0.139)	(0.158)	(0.223)	(0.0002)	(0.0001)				
4	0.004	0.269**	0.261	0.045	0.0002	0.0001	0.005***	-5.618	-5.318	0.893
	(0.005)	(0.133)	(0.159)	(0.253)	(0.0002)	(0.0001)	(0.002)			

Note: Standard errors in parentheses.
*** Significant at 1% level of significance.
** Significant at 5% level of significance.
* Significant at 10% level of significance.

On the other side, table 3.14 below suggests that the equation two is the best specification for the inflation equation. Again, the third lag of inertia is insignificant, but the difference with the AR(3) model is that the changes in food prices are also significant.

Table 3.14. Estimated MA(3) inflation model, Estonia

	Parameter estimates							Diagnostics		
	μ_π	a_1	a_2	a_3	d_1	d_2	d_3	*Akaike i.c.*	*Schwarz i.c.*	*Q(10) p-values*
1	0.014***	0.384***	0.181	-0.286*				-5.544	-5.375	0.720
	(0.005)	(0.129)	(0.129)	(0.154)						
2	0.013***	0.385***	0.217*	-0.242	0.0003*			-5.574	-5.371	0.760
	(0.005)	(0.124)	(0.129)	(0.154)	(0.0001)					
3	0.019***	0.109	0.432***	-0.616***	-0.0007	0.0002*		-5.664	-5.409	0.234
	(0.005)	(0.113)	(0.128)	(0.099)	(0.0001)	(0.00010				
4	-0.003	0.203**	0.062	0.754***	0.0002	0.0001*	0.005***	-5.692	-5.400	0.574
	(0.003)	(0.089)	(0.134	(0.117)	(0.0002)	(0.0007)	(0.001)			

Note: Standard errors in parentheses.
*** Significant at 1% level of significance.
** Significant at 5% level of significance.
* Significant at 10% level of significance.

Based on the information criteria, equation 3 is suggested as the best model fit for the inflation equation. However, without the changes in food prices, the multivariate model for Estonia did not converge, so this variable was included in the signal equation of inflation:

$$\pi_t = a_1\pi_{t-1} + a_2\pi_{t-2} + b_1c_t + b_2c_{t-1} + d_1\Delta FP_t + e_t \qquad e_t \sim iid(o, \sigma^2) \qquad (3.6)$$

In the following table the estimated parameters of the multivariate UC model will be presented.

Table 3.15. The starting and estimated multivariate UC model coefficients, Estonia

Starting parameter values		Estimated coefficients
μ_g	1.6	0.565***
		(0.199)
σ^2_ε	0.226	0.297
		(0.847)
σ^2_v	0.245	0.257
		(1.380)
ϕ_1	1.5	1.743***
		(0.318)
ϕ_2	-0.6	-0.765**
		(0.336)
λ	-0.326	0.058
		(26.252)
μ_π	0.012	0.122
		(114.318)
a_1	0.104	0.240*
		(0.139)
a_2	0.008	0.107
		(0.176)
b_1	-0.006	0.008***
		(0.002)
b_2	-0.006	-0.006***
		(0.001)
d_1	0.0003	0.0003
		(0.0002)
σ^2_π	0.013	0.0002***
		-0.234
Output equation		
H0: No serial correltation:		Q(10) = 0.123
H0: Normality in the residuals:		Jarque-Berra = 0.072
Inflation equation		
H0: No serial correltation:		Q(10) = 0.523
H0: Normality in the residuals:		Jarque-Berra = 0.427

Note: Standard errors in parentheses. The log-likelihood of the function is 110.38. The model converged after 263 iterations.

*** Significant at 1% level of significance.
** Significant at 5% level of significance.
* Significant at 10% level of significance.

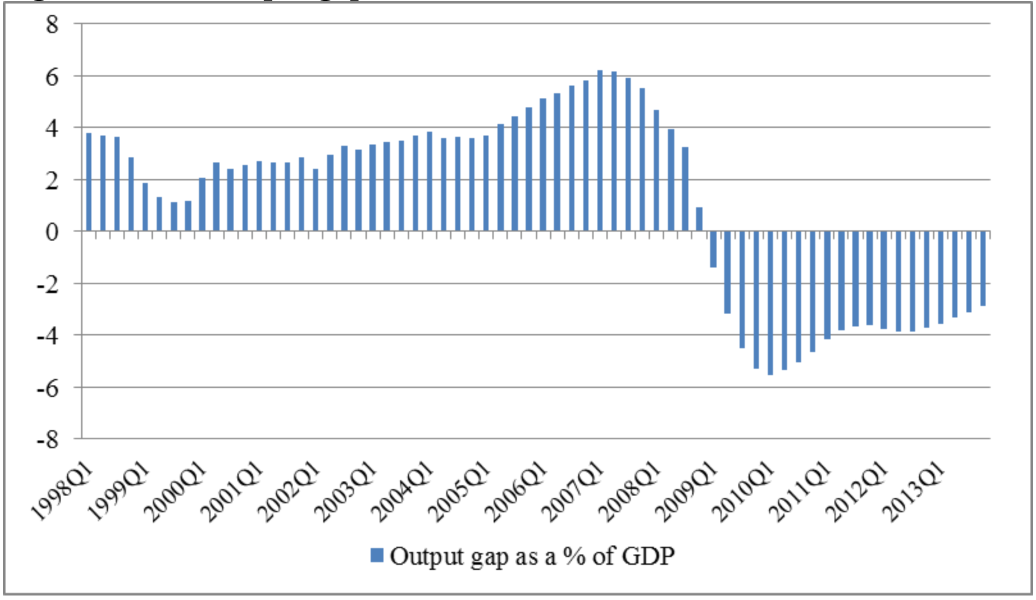
As can be seen in table 3.15, the diagnostics suggest that the multivariate UC model of Estonia is well specified model with regard to the serial correlation and normality of the residuals. Thus, a one pp increase in the inflation rate of the previous period, on average, will increase the inflation rate in the current period by 0.012 pp, ceteris paribus. The second lag is insignificant, meaning that probably expectations do not play a significant role and the market has a short-term memory. From orthodox economic theory, another important link in this model is the relation between the inflation rate and output gap. The output gap of the current period is highly significant indicating that, on average, a one pp increase in the output gap ratio to GDP in the current period will decrease the inflation rate by 0.006 pp, keeping other things constant. Moreover, the previous period's output gap ratio is also highly significant, suggesting that a one pp increase in the previous periods of output gap ratio will also decrease the inflation in the current period by 0.006 pp. Thus, even though the output gap is significant, the economic significance is clearly less so. Nevertheless, the significant current and previous output gap in the model confirm the importance of estimating the output gap for effective policy making in Estonia.

The coefficient of the 2008 dummy variable in the multivariate model was also insignificant, once more emphasising that the crisis had only a transitory impact on Estonia's potential output.[69] The positive coefficient value of the potential output variance probably suggests that the model has found more than one shock, however since the variance is insignificant, the shocks may not have had a permanent impact on potential output path, since the potential output follows a relatively smooth but steady path without significant interruptions (figure 3.4).

Regarding the underlying economic relations in the model, the lagged inflation is statistically significant, which indicates that inertia is an important determinant of the inflation rate in Estonia.

[69] Even though the dummy variable is insignificant in the multivariate model also, we have kept it, not only for informative reasons, but also because the model did not converge even after 500 iterations.

Figure 3.4. The output gap of the multivariate UC model, Estonia



As for the estimated potential output and output gap in figure 3.4, at first sight it appears that they are the same as in the univariate model. However, compared to the univariate case, the differences are a bit larger. For example during the overheating period, the output gap is on average 0.2 pp of GDP larger than in the univariate case, whereas during the recession period the negative output gap is 0.05 pp of GDP larger. The turning point of the cycle is the same in both models that is in Q1 2009.

In the following section, the estimations for potential output and the output gap in Kosovo will be presented.

3.6. The output gap in Kosovo

3.6.1. Unit root tests of the GDP series, Kosovo

The unit root tests for GDP series of Kosovo are reported in table 3.16 below.

H_0: The series is non-stationary (unit root):

Table 3.16. Unit root tests of the univariate UC model, Kosovo*

Country	Series	ADF Test, intercept included	ADF Test, intercept & trend included	Phillip-Perron Test	Perron test with structural breaks	Decision
Kosovo	**Log-levels of quarterly real GDP**	Not rejected at any level of significance.	Not rejected 1%, rejected at 5% level of confidence.	Not rejected at any level of significance.	Not rejected at any level of ignificance.	Non-stationary
	Δ Log GDP	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	-	Stationary

*Note: If the trend in ADF (intercept+trend) test turned out as insignificant, the ADF (intercept only) and Phillip-Perron test were considered to judge a series as stationary or not.

An important property of our model is the presence of the unit root in the real GDP for all three countries, which is confirmed through the ADF tests for non-stationarity. The failure to reject the unit root test in log-levels of GDP series in Kosovo means that potential output will be modelled as a random walk with drift, where the drift term will be modelled as a constant term. Given that we have excluded the 2000 and 2001 data (due to a possible break in the series) and the series starts from Q1 2002, it may be reasonable to assume a constant drift term.

3.6.2. Starting parameter values of the UC model, Kosovo

In short, the initial parameter values set up the same way as for other two countries, are presented in Table 3.17:

Table 3.17. Initial parameter values of the univariate model, Kosovo

μ_g	σ^2_ε	σ^2_v	ϕ_1	ϕ_2
1.5	0.2	2.0	1.4	-0.5

3.6.3. Univariate UC model, Kosovo

After running the univariate regression of the UC model, the estimated parameters of this model are as in table 3.18.

Table 3.18. Estimated parameters of the univariate UC model, Kosovo

Parameter estimates					Diagnostics	
μ_g	σ^2_ε	σ^2_v	ϕ_1	ϕ_2	Serial correlation Q(10)	Normality
0.43***	0.2	2.0	0.510***	0.393***	0.183	0.072

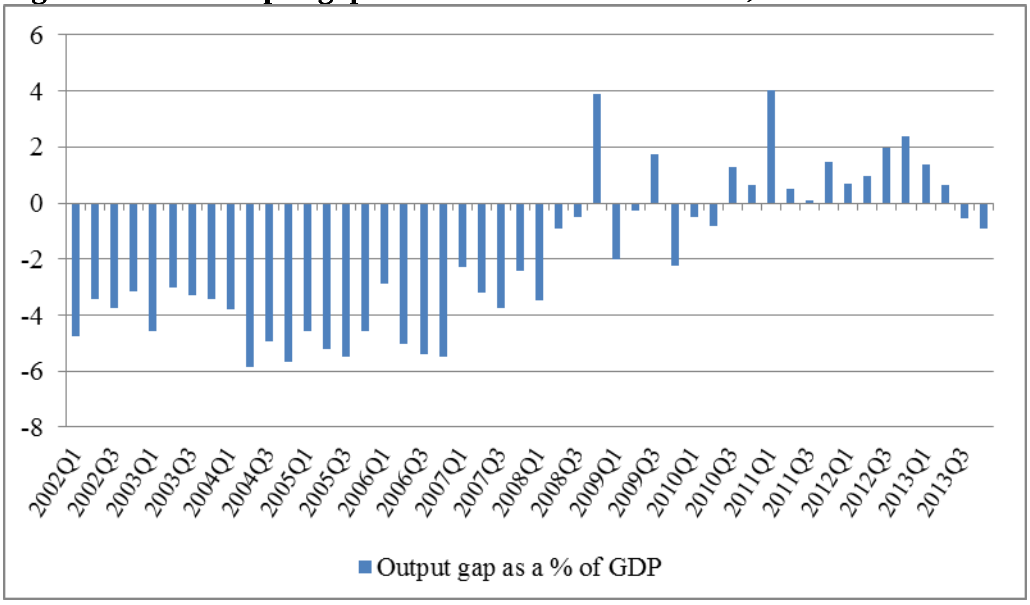
Note: The log-likelihood function of the model is 108.2. The model converged after 20 iterations.
***Significant at 1% level of significance.

As can be noticed in table 3.18, the model is not misspecified with regard to serial correlation and normality of the residuals.[70] The estimated potential output growth rate μ_g is around 0.4 percent per quarter, or 1.6 percent annually, meaning that it is lower than the average growth rate of the actual output (1.5 percent quarterly, see Table 3.15). Also, the potential output growth rate is significant at one percent level, implying that the trend of Kosovo follows a

[70] A one-time dummy was initially added in the model for Q4 2008, in order to control for an structural break in the series caused by Global Financial Crisis, however, the dummy turned out as insignificant (apparently due to slower economic development rather than a crisis or a shock in the economy), thus it was dropped from the model.

random walk with a drift. Both autoregressive coefficients ϕ_1 and ϕ_2 turned out highly significant, confirming that the output gap (cycle) in Kosovo follows an autoregressive process of order two, AR(2). The variance of the potential output and the output gap were calibrated, so they are the same as the initial values in table 3.3[71] turned out highly significant; however the magnitude is not significantly different from zero with no structural breaks identified in the estimated series.

Figure 3.5. The output gap of the univariate UC model, Kosovo



The highest value of the output gap of +4 percent of GDP is identified in the Q1 2011, whereas the lowest value of -5.8 percent is identified in Q2 2004. During the period 2002 to 2008, Kosovo experienced a slow but fairly steady growth rate, however the results suggest that the economy was performing below its potential (Figure 3.5).This was the period when many structural reforms, such as the building-up of the institutional setting, the establishment of a market-based economy, the development of the banking system, the undergoing privatisation process, encouraging trade and attracting investments, etc. were taking place, meaning that the economy did not utilize its resources to their full potential.

In addition, this period was characterised by a relatively low and stable inflation rate of 2.2 percent, on average. The highest historical inflation rate of 9.4 percent occurred at the end

[71] The variances of potential output and output gap had to be calibrated, since the estimated ones were not acceptable for the program and the model did not converge. After trying many combinations of these parameters, the variances of 0.2 and 2 were given to the program, so the EViews did not try to estimate them.

2008, mainly because of the international oil and flour price jumps at the time. The period 2002 – 2008 when Kosovo economy was performing below its potential coincides with a period when the authorities were continuously conducting conservative fiscal policy, and usually concluding the fiscal year with a surplus. This was due to the limited monetary policy tools and the inability of the authorities to finance budget deficit. At the end of 2008 the government starts expansionary fiscal policy, the higher budget spending funded from surpluses in previous periods, led to an almost balanced budget in 2008 and a low budgetary deficit by the end of 2009. The balanced budgetary stance once more coincides with the timing of an eventual balanced state of the economy, i.e. the gap between actual output and the potential output closes somewhere between 2008 and 2009. Additionally to the expansionist fiscal policy, the initial phase of expansion of the business cycle starting from the last quarter of 2008 is also consistent with other economic developments, such as the expansion of bank lending activity. However, on the other side, the indirect impact from the global crisis (decline in exports, slowdown of investment inflows etc.) slowed the pace of economic growth in Kosovo. Further, net foreign direct investments in 2008 and 2009 declined consecutively, indicating that the growth of the potential output at the time may have fallen. Further, a lower level of consumption and especially the decline in prices in international markets led to the lowest inflation rate of -2.4 percent in 2009. It should be noted that the shift of the actual output above potential in two quarters (end of 2008 and beginning of 2009), when the economy is still supposed to be below or almost in equilibrium level, should be considered as an anomaly in the model.

The business cycle of Kosovo experiences a turning point in 2010, where the economy starts performing above its potential output. However, the above potential performance of the economy during Q1 2010 to Q4 2011 is relatively small (0.7 percent of GDP on average) and with variation. The overheating period is consistent with the growing demand pressures driven by private consumption (CBK, 2011) and substantial increases in the public spending (Ministry of Finance, 2011). The increase in domestic demand led to a two-digit increase of imports and an increase in prices (3.5 percent in 2010, 7.5 percent in 2011). Even though the above potential position of the economy is considered as a state of economy where the pace of production capacity cannot cope with the increasing demand, the Kosovo case represents a different case, as even in the below potential phase or equilibrium, most of the domestic demand was fulfilled via imports. Hence, the overheating phase was also reflected in terms of increased imports. Further, the rise in inflation (which in overheating state may frequently be considered as accelerating and unsustainable) is relatively low (one digit). Despite these developments that led to above the potential performance of the economy, on the other side, the lagged effects of the global crisis, mainly via the external sector channels and the uncertainties perceived by the banking sector, may have slowed the pace of growth. However, the overheating period had only a short duration as a consequence of a one-time supply shock, because the actual output soon returned below potential at the beginning of 2012. Moreover, even during 2010-2013 the Kosovo economy was still performing with considerable underutilised resources, as indicated by the relatively high unemployment rate. Therefore, the potential output during this period was quite close to actual output, where the positive output gap comprised only a small proportion of the actual output.

3.6.4. Multivariate UC model, Kosovo

The multivariate model was estimated following the same procedure as in the case of the Czech Republic and Estonia.

The augmented Phillips curve in Kosovo case will be as in the following:

$$y_t = g_t + c_t \qquad (2.1)$$

$$\pi_t = a(L)\,\pi_{t-1} + b(L)D_t + c(L)z_t + e_t \qquad e_t \sim iid(0, \sigma^2) \qquad (3.1)$$

$$g_t = \mu_g + g_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma^2_\varepsilon) \qquad (2.3)$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \upsilon_t \qquad \upsilon_t \sim iid(0, \sigma^2_\upsilon) \qquad (2.10)$$

The inflation equation for Kosovo is as follows:

$$\pi_t = \mu_\pi + a(L)\pi_{t-1} + b_1 c_t + b_2 c_{t-1} + d_1 \Delta FP + d_2 \Delta OP + d_3 \Delta GDP_{t-1} + e \quad e_t \sim iid(0, \sigma^2) \quad (3.4)$$

Following the Box-Jenkins approach, initially we test each exogenous variable for stationarity (table 3.19).

Table 3.19. Unit root tests for inflation variables, Kosovo*

Country	Series	ADF Test, intercept	ADF Test, intercept & trend	Phillip-Perron Test	Decision
Kosovo	**Log CPI**	Not rejected at any level of significance.	Not rejected at any level of significance.	Not rejected at any level of significance.	Non-stationary
	Δ Log CPI	Rejected at all levels of significance.	Not rejected at 1%, rejected at 5% level of significance.	Rejected at all levels of significance.	Possibly Stationary
Exogenous Phillips-Curve Variables (all countries)	**Log Food Prices**	Not rejected at any level of significance.	Not rejected at any level of significance.	Not rejected at any level of significance.	Non-stationary
	Δ Log Food Prices	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	Stationary
	Log Oil Prices	Not rejected at 1% and 5%, rejected 10% level of significance.	Not rejected at 1%, rejected at 5% and 10% levels of significance.	Not rejected at all levels of significance.	Possibly Non-Stationary
	Δ Log Oil Prices	Rejected at all levels of significance.	Rejected at all levels of significance.	Rejected at all levels of significance.	Stationary

*Note: If the trend in ADF (intercept+trend) test turned out as insignificant, the ADF (intercept only) and Phillip-Perron test were considered to judge a series as stationary or not.

As can be noticed in table 3.19, the inflation variables are considered to be integrated of order one.

The residual correlogram of the inflation series was checked, looking for the point where the ACF or PACF fall abruptly. However in the Kosovo case, both, the ACF and the PACF fell abruptly after lag 2, hence we conducted both, an AR(2) model and a MA(2) model, in a pursuit of finding the best inflation model.

Last, based on the ACF and PACF information, we estimated an AR(2) model (table 3.20):

Table 3.20. Estimated AR(2) inflation model, Kosovo

	Parameter estimates						Diagnostics		
	μ_π	a_1	a_2	d_1	d_2	d_3	*Akaike i.c.*	*Schwarz i.c.*	*Q(10) p-values*
1	0.879 (0.922)	0.401** (0.192)	-0.050 (0.448)				4.965	5.150	0.756
2	0.117 (0.252)	0.148* (0.078)	-0.107 (0.073)	0.524*** (0.035)			2.706	2.937	0.833
3	0.116 (0.257)	0.148* (0.079)	-0.106 (0.076)	0.523 (0.040)	0.0002*** (0.006)		2.770	3.048	0.832
4	0.132 (0.252)	0.147* (0.081)	-0.104 (0.078)	0.525*** (0.042)	-0.0003 (0.006)	-0.017 (0.046)	2.829	3.153	0.831

Note: Standard errors in parentheses.
*** Significant at 1% level of significance.
** Significant at 5% level of significance.
* Significant at 10% level of significance.

Based on statistical validity, statistical significance and the smallest Akaike and Schwarz information criteria, equation 2 is the best fit model for the inflation equation.

Table 3.21. Estimated MA(2) inflation model, Kosovo

	Parameter estimates						Diagnostics		
	μ_π	a_1	a_2	d_1	d_2	d_3	*Akaike i.c.*	*Schwarz i.c.*	*Q(10) p-values*
1	0.106 (0.721)	0.683*** (0.149)	-0.370*** (0.094)				4.518	4.699	0.770
2	0.964 (0.283)	0.199** (0.061)	-0.182** (0.058)	0.511*** (0.036)			2.542	2.769	0.553
3	0.081 (0.287)	0.205*** (0.0621)	-0.181*** (0.057)	0.503*** (0.039)	0.003 (0.004)		2.587	2.859	0.641
4	0.100 (0.298)	0.205*** (0.063)	-0.181*** (0.059)	0.502*** (0.405)	0.002 (0.005)	-0.006 (0.018)	2.643	2.961	0.597

Note: Standard errors in parentheses.
*** Significant at 1% level of significance.
** Significant at 5% level of significance.
* Significant at 10% level of significance.

Consistent with the AR(2), the drift term is also statistically insignificant in the MA(2) equations (Table 3.21). Contrary to the AR(2) specification, both autoregressive coefficients of inertia turned out as highly significant. The changes in food prices are significant in all specifications. However, changes in oil prices and lagged growth rate of the GDP turned out as insignificant. The information criteria suggest that the equation 2 is the best model fit for the inflation equation in Kosovo.

In choosing between the AR(2) and MA(2) model, the Akaike and Schwarz suggest that the MA(2) model is a better specification, so we proceed using the following Phillips curve equation for Kosovo:

$$\pi_t = \mu_\pi + a_1\pi_{t-1} + a_2\pi_{t-2} + b_1c_t + b_2c_{t-1} + d_1\Delta FP_t + e_t \qquad e_t \sim iid(o, \sigma^2) \qquad (3.2)$$

After specifying the inflation equation, as discussed in the univariate unobserved model, we also need to set initial parameter values for the multivariate model. The initial parameter values for the drift term μ_π, autoregressive coefficients ϕ_1, and ϕ_2 of the output gap and the variances of the potential output and the output gap, σ^2_ε and σ^2_υ, are set in the same way as in the univariate model. With regard to the inflation equation initial parameter, we have tried to estimate the above specified equation (3.2) using linear regression, however the coefficients of the same equation but using the MA(2) specification yielded better parameter estimates.
The parameter estimates of the multivariate UC model will be presented in table 3.22.

Table 3.22. The starting and estimated parameters for the multivariate model, Kosovo

Starting parameter values		Estimated coefficients
μ_g	1.5	0.4*** (0.075)
σ^2_ε	0.2	0.2 (0.072)
σ^2_v	2	2 (0.09)
ϕ_1	1.4	0.51*** (0.079)
ϕ_2	-0.5	0.393*** (0.076)
a_1	0.174	0.105 (0.079)
a_2	-0.145	-0.064 (0.076)
b_1	0.039	-0.057 (0.113)
b_2	-0.058	-0.039 (0.103)
d_1	0.519	0.534*** (0.029)
σ^2_π	0.795	-0.408
Output equation		
H0: No serial correltation:		Q(10) = 0.183
H0: Normality in the residuals:		Jarque-Berra = 0.072
Inflation equation		
H0: No serial correltation:		Q(10) = 0.814
H0: Normality in the residuals:		Jarque-Berra = 0.251

Note: Standard errors in parentheses. The log-likelihood of the function is 148.28. The model converged after 46 iterations.

*** Significant at 1% level of significance.

** Significant at 5% level of significance.

* Significant at 10% level of significance.

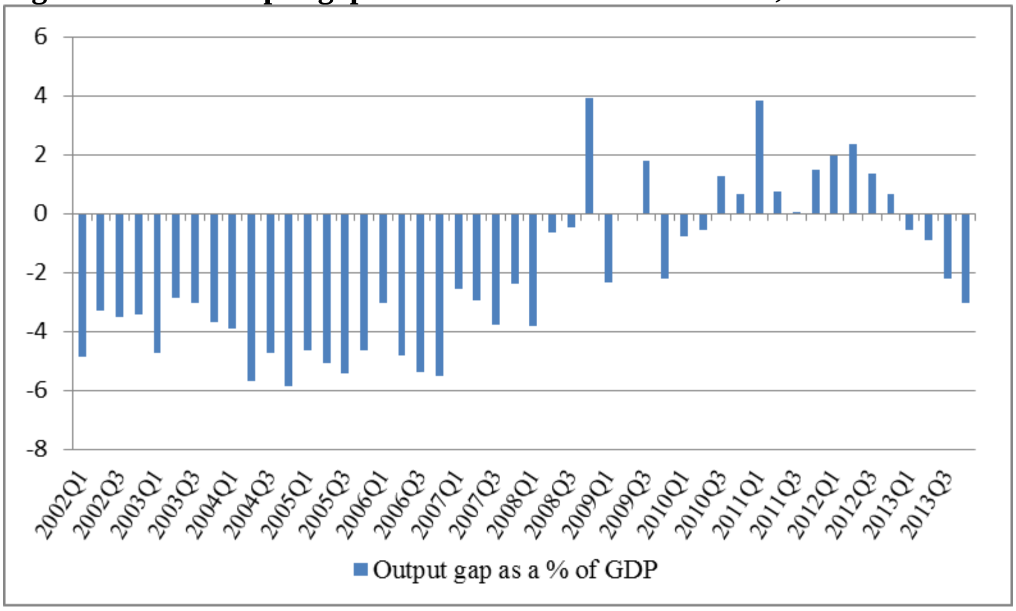
The diagnostic tests suggest that the multivariate UC model for Kosovo is well specified. The residuals of the model, similar to the univariate model, are serially independent and normally distributed. The relation of the output gap and inflation in the case of Kosovo is insignificant, however the negative sign is theoretically suggestive. An interpretation of this result may be that when the economy is performing below the potential output, it is easier to activate the work force at the same price level. Meanwhile, if the aggregate demand increases, the need for extending productivity increases, too, however the employees may also require higher payment. This in turn may increase the price level. However, the insignificant relationship between output gap and the inflation rate may also reflect the specific labour market conditions in Kosovo, where even though the unemployment rate is considerably high, the reservation wage may also be relatively high due to the effect of remittances, whereas wages are not necessarily flexible in the short-run. Therefore, the impact of the output gap on the inflation rate may be insignificant. Additionally, despite the presence of a positive or negative output gap, as argued in section 2, Kosovo has a very limited scope to use monetary policy to influence the prices.

In line with expectations, food inflation turned out to be an important determinant of the domestic CPI inflation; that is, a one pp increase in food inflation, on average leads to 0.5 pp increase in the overall CPI, ceteris paribus, on a quarterly basis. This finding is expected in the case of Kosovo, given that the foodstuff represents around 60 percent of the CPI basket.

Further, the potential output growth rate is 0.4 percent quarterly, which is lower than the actual output growth rate. Both autoregressive coefficients turned out as highly significant, confirming that the output gap follows and AR(2) process, as in the univariate model.

Finally, an examination of potential output and the output gap series presented in figure 3.6 will follow.

Figure 3.6. The output gap of the multivariate UC model, Kosovo



The fact that the lagged terms of the output gap in Kosovo turned out to have an insignificant impact in inflationary pressures, was reflected in the generated unobserved series. As can be noticed the multivariate model did not turn to be more informative than the univariate model since the estimated UC results are only slightly different from the univariate specification. In most of the quarters, the output gap is the same, in some others the output gap differs only in the decimals. For example, during 2006, the output gap in the multivariate model is larger for 0.00017 pp, whereas during 2010 and 2011 by only 0.0002 pp of the GDP.

The period between 2008 and 2009 represents an important transition of the business cycle in Kosovo. From 2008 onwards, the Kosovo business cycle experiences two turning points: the first being between 2008 and 2009 where the economy, after around ten years performing below its potential reaches 'equilibrium' and the second turning point during 2010, when the economy starts performing above its potential output, even though the potential output is quite close to the actual output. To conclude, the business cycle in Kosovo, despite other important developments in other sectors of the economy, throughout the period under examination closely mimics fiscal policy developments. Therefore, fiscal policy in Kosovo economy appears to have played a predominant role in the business cycle of Kosovo. Overall, developments like the declaration of independence in 2008 and later the moderate effect of the global financial crisis had only a transitory impact in the potential output (trend), whereas the output gap (cycle) turned from negative into a positive one. This is in line with the main assumption of the UC model, indicating that shocks affecting trend and cycle are not correlated.

3.7. Conclusion

After elaborating the theoretical part of the business cycle in the previous chapter, in this chapter we have turned to estimating potential output and the output gap. We have presented estimations of the univariate and multivariate unobserved components model, using three representative countries of the European Transition economies, namely the Czech Republic, Estonia and Kosovo.

The univariate model, considered to be more statistical in nature as it only exploits the real GDP series to extract the unobserved components, was expected to deliver relatively inferior results, compared to the multivariate model. This is because in the multivariate model we have also utilised orthodox economic theory via the Phillips curve. This has proven to be true in the case of the multivariate UC for the Czech Republic, as the estimated potential output and output gap were more informative and corresponding to the economic situation. As suggested by the theory, a negative relationship between inflation and output gap in the Phillips curve equation was established in the case of all three countries. Even though the relationship between inflation and output gap turned out as insignificant, the multivariate specification for the Czech Republic was informative also because it suggested a presence of inflation inertia in the country. In the multivariate specification, inertia also turned out to be an important determinant of inflation rate in the Estonian economy. The findings in the Czech Republic regarding the Phillips curve relationship were also repeated in the case of Estonia. In Estonia the results suggested that current and previous period output gap influence the level of inflation and may be important in conducting monetary policy in the country, confirming the importance of the output gap in the policy making. However, in the case of Kosovo the results did not change much in the multivariate model, nevertheless this model was useful in suggesting that the output gap is probably not useful in decision making (due to the lack of the monetary policy instrument), as it was found that fiscal policy plays the predominant role in the business cycle.

The generated results from both, univariate and multivariate generated output gaps within intervals of -6 to +6 percent of GDP for the three selected countries. Based on the theory of business cycle in transition economies these represent reasonable results (Darvas and Vadas, 2003). The impact of the Global Financial Crisis controlled for in the univariate model, turned out to be insignificant, suggesting only a transitory impact of the crisis in the potential output for all three countries. However, in the multivariate specification the impact of the crisis turned out to have a significant impact on the potential output of the Czech Republic.

Overall, as discussed in chapter 2, multivariate approaches have a significant advantage over the univariate ones, since they incorporate additional economic information in the estimation of potential output and the output gap, thus potentially increasing the accuracy of the estimates. At the same time, multivariate methods are far more difficult to implement in practice, which makes them less popular than univariate methods. Either way, the estimates of the potential output and output gap incorporate a considerable amount of uncertainty, and none of the methods used yield ideal results. Thus, good practice in this context is to use several univariate and multivariate techniques' estimates and then compare them. As Grewal and Andrews (2011) points out, the multivariate techniques exploit structural information to extract the unobserved variables, thus tend to be the most accurate. Yet, large systems do not necessarily outperform small systems, since although the former may increase accuracy they also create more uncertainty (more chances of misspecification error).To summarise, we may conclude that the multivariate UC model may bring additional useful information regarding the potential and output gap; however in our cases the multivariate model did not always outperformed the univariate model. Hence multivariate UC models bring additional information into the modelling but also appear to increase the complexity of the procedure and uncertainty of the findings.

After investigating the definition and behaviour of the output gap in the previous and current chapter, in the following chapter we will examine the definition and determinants of excess liquidity in European transition economies. The relationship between output gap and excess liquidity will follow in chapter 5.

Chapter 4. The determinants of excess liquidity in European transition countries

4.1. Introduction

The central part of this chapter will be devoted to defining and measuring excess liquidity in European transition countries. As will be seen in the upcoming section, the definition of excess liquidity in the context of the transition countries is not a straightforward matter. While the concept of excess liquidity sounds intuitive, its measurement is fraught with difficulties. First of all, different studies tend to define and measure excess liquidity differently. Second, as argued in chapter 1 (banking structure and growth levels in these countries), the relevance and importance of excess liquidity is particularly linked to the overall state of the economy of the European transition countries. Thus, while other studies usually treat excess liquidity in the light of the effectiveness of monetary policy, we emphasise the importance of excess liquidity as a one of the most important but yet unutilized resources in the domestic economies of these countries. Having said that, a fundamental concern of this research programme is whether the persistent presence of excess liquidity in many European transition countries is impeding growth and widening the output gap, as discussed in chapter 1. From the transition countries context, we will try to theoretically contribute to the definition of this concept in relation to economic growth. After explaining the preferred definition of excess liquidity, the analysis explores its pattern and potential determinants over time and across countries. In addition the discussion in this chapter will add to the limited theoretical and empirical analysis of the role of excess liquidity as an impediment to growth.

What exactly is liquidity, and how do we determine whether it is excessive or not? Why do banks seemingly hold excess liquidity in face of more profitable lending opportunities? Is it a demand side or supply side problem? These are the main research questions that we address in the upcoming sections.

The rest of the chapter is organised as follows. Section 4.2 revisits the previous definitions of excess liquidity, proposes a new definition and elaborates the theory on determinants of excess liquidity. Section 4.3 explains the methodology used to calculate excess liquidity based on the new definition proposed, the model specification as well as a discussion on the estimated results. Section 4.4 presents the disaggregated portion of precautionary versus excess liquidity. Section 4.5 concludes the chapter.

4.2. Theoretical background

4.2.1. Definition of excess liquidity

In mature market economies, financial sectors are typically well developed, with both effective primary markets with a large number of players and active secondary markets where financial securities can be traded with ease. In these economies, the monetary authorities typically rely upon open market operations in securities, either government issued or their own papers of different maturities, to clear any surplus funds (Ruffer and Stracca, 2006). Hence, with the exception of the recent past after the onset of the Global Financial Crisis (GFC), the incidence of 'excess liquidity' was relatively rare in developed countries and it appears more as a phenomenon characterising developing and transition economies. As a consequence, the concept of *excess liquidity* or *excess reserves* is frequently treated imprecisely, if at all, in the literature of developed countries, though it has also been neglected in empirical studies for developing/transition countries.

Conceptually, liquid assets represent the sum of the financial assets that can be quickly mobilized for lending or other purpose transactions. However, the question posed is how much of those liquid assets can be considered to be in excess of what is needed for banks to conduct daily transactions in the intermediation process and/or comply with the regulations of the domestic money market regulators, i.e. which part is not considered as 'excess'. Different authors, based on theory and the practices of central banks, define excess liquidity in different manners.[72] For example, at the global level, Ruffer and Starcca (2006) define excess liquidity as the inverse of the 'velocity of the money', indicating that part of the money supply is not circulating in the economy but being kept in the banking system. At the bank level, Saxegaard (2006) who studied the causes of excess liquidity in two Central African countries, Nigeria and Uganda, argues excess liquidity is typically equated to the quantity of cash and deposits of commercial banks held with the central bank in excess of the required or statutory level. Here, the excess liquidity ratio is calculated as a percentage of the level of excess liquidity to total deposits (demand and time deposits, savings and foreign currency deposits). Other authors that examine excess liquidity and its determinants in a similar fashion in transition countries include Agenoret et al., (2004), Aikaeli (2006) and Jayaraman and Choong (2012).

[72] We should note that the definition of excess liquidity is closely related to its preferred measurement methodology/formula, because the conceptual definition determines its measurement.

Excess liquidity thus measured may be presented via the expression presented below:

$$EL = d - RR \tag{4.1}$$

$$ELR = EL / D \tag{4.2}$$

In equation (4.1) *EL* represents excess liquidity in levels, *d* are deposits of commercial banks held at the central bank in levels, RR are required reserves from the regulator, *D* are total commercial banks deposits and *ELR* is the excess liquidity ratio.

Another set of authors define excess liquidity as the total commercial bank liquid assets minus the required liquid assets determined by the central bank (Khemrraj, 2006, 2007). In contrast to the first definition of excess liquidity, this means that other parts of liquid assets may also be considered as excessive. Heenan (2005) supports this argument by stating that excess liquidity may include other assets such as treasury bills, which can be quickly converted into commercial banks' loans.[73] The second definition may be expressed as follows:

$$EL = \sum LA - RR \tag{4.3}$$

We should note that in total liquid assets *LA* are:

$$\sum LA = \sum Assets - Loans - Fixed\ Assets - Other\ Non\text{-}Liquid\ Assets = Cash + Balance\ with\ CB + Placements\ Abroad + Securities^{74} \tag{4.4}$$

$$\Rightarrow EL = (Cash + Balance\ with\ CB + Placements\ Abroad + Securities) - RR \tag{4.5}$$

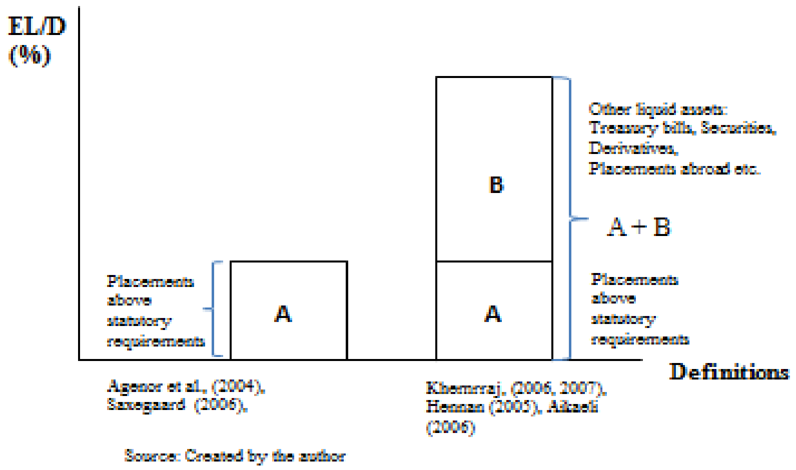
We can agree with the first definition in equation (4.4), because it comprises part of excess liquidity. However, the other part of excess liquidity measured as in equation (4.5) is considered as imprecise. This is because, the entire portion of excess minus required reserves cannot be considered as excess neither from central bank's perspective, nor from the commercial banks' point of view. This is because, even if banks do not have an explicit mandatory liquidity ratio set by the regulator (e.g. liquid assets to short-term liabilities or

[73] Rather than being concerned how quickly these assets convert into commercial banks' loans we are primarily interested in knowing what is the level of excess liquid assets that is kept in all forms rather than being lent in the domestic economy and later on what is the impact of the excess liquidity in these economies.

[74] Placements and deposits abroad are usually instruments invested with maturity date up to 30 months and less than 1 year, hence on average they are considered as liquid assets. Marketable securities are very liquid as they tend to have maturities of less than one year. Furthermore, the rate at which these securities can be bought or sold has little effect on their prices.

liquid assets to total assets should be held at all times at 20 percent), some other liquidity restrictions usually hold, be it structure of liquid assets, maturity gaps, internal requirements to be set by the banks' management, deposit concentration, stress-test findings and limitations, etc. Hence, a part of all liquidity assets as defined in equation (4.5) will be restricted by regulatory requirement and cannot be all considered to be 'excess'. Instead, we propose a new definition of excess liquidity as proposed in figure 4.1, whereas its measurement formula will be presented in more detail in section 4.3.

Figure 4.1. The conceptual definition of excess liquidity



A = Placements above statutory requirements at the central bank

B = Other excess liquid assets above mandatory liquidity ratio (Treasury bills, securities, derivatives, placements abroad etc.)

Area A represents the part including only the deposits above statutory requirements held at the central bank, whereas area B depicts other liquid assets that are held above mandatory ratios set by the regulator (see section 4.3.1 for calculations). These assets may be assets held in commercial banks' accounts such cash, invested in treasury bills, or even investments abroad such as securities and placements/deposits in other banks abroad. Based on the first group of definitions, excess liquidity would quantitatively be equal to area A, whereas based on the second definition, the level of excess liquidity is much higher considering other liquid assets and equals areas A+B.[75]

After revisiting previous definitions of excess liquidity and proposing our preferred method of measuring it should be noted that all of the excess liquidity thus measured may not be necessarily considered as 'excess' . This is because, given that banks are profit maximising firms, 'excess' liquidity accumulation (or a part of it) may simply reflect an optimising decision. A decision whether excess liquidity may be considered as 'excess' in the meaning of being an unutilised resource for further investments can be taken only after disaggregation between the precautionary part and involuntary part of excess liquidity. The precautionary part of excess liquidity is considered to be held voluntarily by the banks, as part of a profit maximising decision, thus this is not treated as excess in our thesis (see section 4.4 for results). In section 4.2.2 sources of excess liquidity will be elaborated and the parts that may be treated as excess from the economy's perspective will be discussed.

Finally, when considering which definition may be more appropriate for our study, we should note that our investigation of excess liquidity does not stop at the identification stage nor are we concerned, as in many other studies, with its impact on the efficiency of monetary policy. Our investigation of excess liquidity is motivated by the discussion in chapter 1, where we argued that many banks in European transition countries while lending low levels of credit compared to developed countries also hold relatively high levels of unutilised resources. Thus, excess liquidity in this study is treated as important unutilized resources that could be translated into further lending and potentially augment growth. Having said that, the second definition considers a larger range of assets that are not invested in the domestic economies, hence this definition may be more appropriate in our case.

4.2.2. Sources of excess liquidity

When arguing about the relevance of excess liquidity, one might pose the question: why is excess liquidity a problem? As argued in chapter 1, banks in European transition economies have continuously accumulated liquid assets above their statutory requirement in central bank accounts that are usually non-remunerated, even in the more developed transition countries

[75] The excess liquidity measured this way will be without any prejudice treated as excess for definition purposes. Whether or not it is excess from the economy's perspective is analysed in sections 4.3 and 4.4.

such as the Czech Republic, Poland, and Hungary. In the cases when banks gain a remuneration rate, which is the case in some of the South-Eastern European (SEE) countries, the rate is typically much lower than the market rates, thus not providing an attractive rate of return.[76]A portion of other liquid assets invested in Treasury bills or placements abroad that comprise a large part, (often more than one of third of total assets, as noted in chapter 1) are usually less risky than investments in loans, as well yielding higher rates of return than deposits at the central bank. Nonetheless these returns are fairly low and in principle do not provide much incentive for banks to hold excess liquidity.[77]

Second, loans are one of the most profitable assets to hold for the commercial banks in these countries, given that the average lending rate is around 11.4 percent, compared to around 4.5 percent lending rate in Euro area countries (see section 1.5). On the other side, the non-performing loan rates in European transition countries (2000 - 2012 average was 7.9 percent), higher than those in the Euro area (2000 - 2012 average was 3.5 percent), (World Bank Data, 2013). However, even after adjusting the return rates with the default rates, banks in transition countries seem to be earning a higher margin on their loans. Yet, the ratio of lending to the private sector/GDP in European transition countries is considerably lower than that in the Euroarea, which suggests deficient lending in transition countries (see section 1.7).[78] It should be noted that, if all of the excess reserves are held for precautionary reasons from banks, be it for conducting daily transactions, than this part of excess liquidity is not necessarily ‘excess’ from the perspective of the banks and the economy, as well. However, if excess reserves are held by banks involuntarily when they would in otherwise invest it (different circumstances), then this part (or all of it) is considered as ‘excess’. In the following, sources of excess liquidity will be elaborated.

[76] The IMF (2005) indicates that in order to encourage banks to trade with each other, the remuneration rate on deposits in central bank should be lower than the cost of borrowing from the central bank.

[77] Among the most common of benchmark interest rate indexes used are London Interbank Offered Rate (LIBOR), Euro Over Night Index Average (EONIA) and Euro Interbank Offered Rate (Euribor), which are used as benchmark rates from banks in Europe and U.S.A. to set up interest rates for placements from other banks. Banks of European transition economies, usually place their deposits based on other banks and gain a return indexed mainly to these two indexes. For example, the average 12-month LIBOR rate during 2000 – 2008 was 3.60 percent and from 2009 – 2014 decreased to 1.04 percent. The average EONIA daily interest rate for 2005 – 2008 was 3.16 percent and from 2009 - 2014 it was 0.40 percent. The average EURIBOR 12-month interest rate during 2000 – 2008 was 3.28 percent and from 2009 – 2014 it was 1.45 percent. The difference between average lending interest rate of 11.4 percent and the average of these indexes of 1.95 percent is approximately 10pp. To conclude, the gap between two rates may lead us to believe that investing in placements abroad provides much less incentive in terms of rate of return as compared to investments in loans.

[78] Based on the World Bank data, the average total domestic credit to GDP in European transition countries in the period 1998 - 2012 was approximately 47 percent, compared to around 112 percent in the Euroarea. Based on these data, the gap of 65pp can be considered as relatively large.

Overall, while banks in European transition countries are holding non-remunerative assets they continue to accept publics' deposits which are also costly to hold (deposit rates around 7 percent). Excess liquidity thus held, seemingly reduces bank's earnings and thereby reduces their capital and increases their vulnerability. On the other side, in a transition country where enterprises may crave for working capital, underutilization of financial resources is likely to be a crucial economic problem (Aikaeli, 2006; see section 1.3). Hence, excess reserves seem to be a scarce resource that may neither yield sufficient returns for banks nor adequate loans for the economy. Therefore, it is in the hand of the banks to choose between constraining finance's role in growth or augmenting that role (Khemrraj, 2007). The obvious question that arises is why do profit-maximising banks accumulate and maintain excess liquidity and continue to accept public's deposits rather than making additional loans?[79] Second, if the excess liquidity were to be lent in the domestic economy, would it raise investment and aggregate demand and hence augment growth in these economies? Or, is the accumulation of excess liquidity actually a consequence of the depressed economies in these countries and lack of demand and low absorption capacity?

The set of countries to be examined in this research includes late-comers to the transition process, aid dependent and remittance receivers such as Albania, Bosnia and Herzegovina. Kosovo and Montenegro, some of the more developed transition countries such as the Czech Republic and Poland that have been catching-up faster with the developed European countries. Thus, it may be difficult to identify a single cause behind the build-up of excess reserves.

In the search for causes behind the accumulation of excess reserves and unlike other studies, such as Agenor et al. (2004), Henaan (2005), Saxegaard (2006) or Khemrraj (2006 and 2007) mentioned earlier, we will extend the discussion of excess liquidity from the perspective of the main macroeconomic theories: Keynesianism and (New) Classical analysis, because these theories indicate much broader points of departure for analysis than just particular theories. They are competing ways of looking at the (macroeconomic) world with competing assumptions which, in turn, inform rather different theories. Other studies even though they

[79] Accepting deposits meanwhile the bank has accumulated excess reserves may also be a long-term and strategic plan to defend and/or built market share. However, only after the disaggregation between precautionary and involuntary excess liquidity will we be able to judge if this is the case.

raise the fundamental questions, i.e. is it a demand or supply side issue, do not approach excess liquidity from these mainstream theories.

Being the main source of finance in transition economies, banks are often being accused of holding excess liquidity, restricting loans and consequently causing depressed economies. Conversely, banks point to insufficient demand for loans, market and institutional inefficiencies so banks claim that they are actually the victims. Alternatively, excess liquidity and economic growth may not necessarily be in a causal relationship. Otherwise, relationship may only represent a consequence of a wider system that determines both: excess liquidity in banks and a lack of demand for loans. In other words, both, excess liquidity and lack of demand for loans, may be outcomes of other factors (deficient demand and low economic growth) either simultaneously or from one period to another.

Based on the Keynesianism view, excess liquidity would be regarded from a wider economic view, i.e. from the level of aggregate demand that creates equilibrium between investment and saving plans. Based on this theory, excess liquidity can be viewed either as a temporary deviation from the equilibrium, or as a consequence of a systemic market failure to adjust (De Vroey and Malgrange, 2011). For example, the banking market could fall prey to inter-temporal coordination problems, such as the inability of the interest rate to clear savings and investment plans, which then were manifested in a build-up of excess liquidity. If this is the case, then one possible outcome may be that these countries may be in a liquidity trap. In effect, assuming that treasury bills and reserves are perfect substitutes, but the lending rate is too low to cover the intermediation costs. Here expectations are that interest rates must rise and bill/bond prices fall, hence fear of capital losses dissuades purchasing securities. In this case, banks would be better off in holding reserves as they would have a higher yield rather than lending them. Hence, a monetary expansion by the central bank just leads to an increase in excess reserves, even beyond banks' prudential requirements, since interest rates cannot be lowered, because open market operations are ineffective Therefore,, an expansionary monetary policy in this case would only encourage banks to build-up even higher levels of excess liquidity. (However, in the case of transition economies we are discussing banking systems with high interest margins, but where banks still hold excess liquidity. Hence, the liquidity trap argument does not appear relevant to the banking system of these countries in recent decades. The savings of the public in the banking system and the investments in loans from the banks may be considered to be in equilibrium at any level of excess liquidity accumulated; however this would create an equilibrium in which not all resources are utilised. From this perspective, excess liquidity in the banking system is consistent with a macroeconomic equilibrium; nonetheless this would represent equilibrium in a depressed economy.

If excess liquidity is a consequence of excess deposits, i.e. the public wants to save more than the banks are willing or able to invest, then this represents a macro level problem. If households and enterprises want to deposit more than they are willing to invest, then the paradox of thrift (savings) may arise: the desired savings can rise above the desired investments. Hence, the increase in savings would lead to a lower level of banks' investment and therefore to a lower aggregate demand and employment, consequently to an even lower level of investment. In this case, if previous savings persist and GDP falls then the excess liquidity is expected to rise. And so on, even though savings and investments may reach multiple equilibria (with different levels of excess liquidity left as a consequence of their mismatch, i.e. equilibrium at a depressed level of economic activity), the new equilibrium would also mean a lower level of economic activity. However, if excess liquidity arises from 'excess' deposits, then why do not depositors find projects to invest in themselves? Maybe this situation represents a signal of the lack of profitable projects to invest, i.e. a lack of absorption capacity.

The Neo-Classical position regarding excess liquidity would be that, just as unemployment is evidence of a malfunction in labour markets, then excess liquidity is an evidence of malfunction in the money and capital markets. Thus, it would be useless for the central banks to interfere in the banking system to reduce the excess liquidity below what may be an equilibrium level. Any central bank interventions via monetary policy may have short term effects but without sustainable long term impact, as the agents in the market, that is depositors and banks optimise their positions (De Vroey and Malgrange, 2011). Thus, corresponding solutions would be found in the capital markets, banking systems respectively, i.e. to strengthen banking competition and eliminate central bank restrictions.

Even though it may be useful to examine the causes of excess liquidity from the Keynesian and Neo-Classical approaches, typically these two approaches assume highly efficient financial systems, referring to the state of the economies in developed countries. In search of a more appropriate answer with regard to transition countries, below we will continue exploring demand and supply side factors having in mind the structure of the economies and banking systems in transition countries discussed in chapter 1. Recent studies that have examined potential determinants of excess liquidity in transition countries start with the question of whether it is a demand or supply side phenomenon.

Firstly, Agenor et al., (2004) study the fall in credit in crisis-stricken East Asia. Their methodology consists of estimating a model of banks' demand for excess liquidity which includes explicitly the precautionary motive for holding excess reserves. The portion of excess liquidity which is involuntary can then be calculated as the difference between statutory excess liquidity and the level of excess liquidity predicted by the model of banks' demand for excess reserves. The authors find that the fall in credit in the aftermath of the crisis and the persistent increase in excess reserves holdings by commercial banks was largely a supply-side phenomenon, as banks reduced lending due the higher risk of default perceived. Their results suggest that the build-up of excess reserves was not excessive in the sense that it did not exceed the desired level of reserves holdings, as it was part of the maximizing profits and cost minimizing behaviour of banks.

Saxegaard (2006) examined the determinants of excess liquidity in Sub-Saharan African countries and found that the increase in excess liquidity in the Sub-Saharan region over the course of the sample period was largely due to an increase in involuntary excess liquidity, i.e. a demand side problem. He argues, that in order to be able to make inferences regarding the implications of excess liquidity, one needs to differentiate also the part of the excess liquidity held voluntarily by the banks.[80] According to the author, the precautionary motive for holding excess liquidity is predictable, in sense that commercial banks willingly would want to hold idle liquid assets unused, whereas involuntary excess liquid assets are constrained by demand side factors. In addition, the author explains that holding precautionary excess liquidity requires banks to be risk averse, whereas banks involuntarily holding excess liquidity may even be risk-neutral. Nevertheless, commercial banks already hold a part of their deposits, as the required rate from the regulatory that should serve as a precautionary measure against liquidity or other related risks, even though the regulators choice of the reserve ratio may not necessarily coincide with the bank's desired precautionary rate. If required reserves serve as a

[80] The consequences of excess liquidity in Saxegaard's (2006) study are important for monetary policy conduct. If excess liquidity is kept for precautionary purposes, expansionary monetary policy may be ineffective and only make banks to hold even more excess reserves. If excess liquidity is kept involuntary in banks, due to for example lack of demand for loans, then an expansionary fiscal policy may help to induce higher aggregate demand and increase the demand for loans, therefore encouraging banks to reduce excessive funds. However, the implications of excess liquidity with regard to monetary policy are outside the scope of this research project.

prudential tool to safeguard banks from unforeseen liquidity contingencies, why do banks see it reasonable to hold more than that? Do banks have more information and better assess the risk than the regulator?

In countries where the required rate of reserves tends to be frequently changed by the central bank, commercial banks may hold additional liquid assets in order to anticipate increases in the required rate. The required reserve rate in European transition countries was higher during the second half of 1990s than at the beginning of 2000s. During the period 2002 to 2014, the reserve required rate followed a decreasing trend in almost all of these countries. The more developed transition countries such as the Czech Republic, Poland and Hungary on average during 2000s had a reserve requirement rate spanning from 9 percent to most recently 2 percent. In our case this argument may not relate to the case of Albania and Kosovo, where the required rate was constant throughout 2000s.[81]

However, even if banks are unable to extend lending, one would except that they can still decrease the level of deposits accepted. In some cases, even when banks are willing to extend lending and reduce excess reserves, they may not be able to do so due to a lack of competition and asymmetric information in the market. In line with this argument, Khemrraj (2006) explains that excess liquidity provides important insights into the loan market in less developed countries. By being the main provider of financial resources in the economy, banks in less developed countries can exert oligopoly power by which they require a *'minimum loan rate'*. Before issuing an additional loan banks require the minimum loan rate that is above transaction costs, market risks and an exogenous reference rate, which makes this rate higher than what would it be under greater banking competition. If the new borrower is unwilling to pay the minimum interest rate, then banks may passively accumulate excess reserves This argument may be in line with the current banking structure in European transition countries, examined in chapter 1, with a relatively small number of banks operating in these countries and especially the high concentration rate in the market of three largest banks.[82] Khemrraj (2006) explains that in Guyana, commercial banks cannot invest their entire holdings of non-

[81] In Albania and Kosovo the required rate did not change from the beginning of 2000s, that is 10 percent of total deposits.
[82] In European transition countries, the number of the banks operating may vary from 9 and 10 in Kosovo and Montenegro to around 45 and 71 in Czech Republic and Poland, respectively. Regarding the market concentration, three largest banks in all of these countries consistently dominate around 70 percent of total banking assets (approximately for the last 15 years), varying from around 55 percent in Bosnia and Herzegovina and Serbia to around 95 percent in Estonia (FRED Economic Data, 2013).

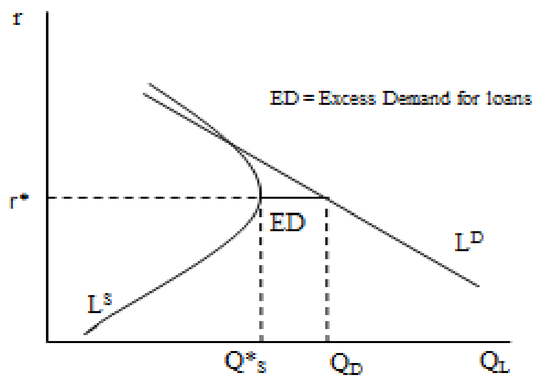
remunerative excess reserves in foreign assets, due to foreign currency constraints set by the central bank. However, to our knowledge, there are no legal barriers to commercial banks in European transition countries investing abroad; indeed investments abroad either in the form of Treasury bills in foreign countries, securities or deposits abroad actually comprise the largest share of excess liquid assets. Domaç et al. (1999) analysed the credit crunch in Asian countries arguing that a widening of the spread between the lending and risk-free asset (e.g. Treasury bill) rates while at the same time credit is falling, represents evidence that the demand for loans could not have fallen more than the supply for loans. In other words, higher spreads would normally encourage banks to lend more not less, meaning that this represents a supply side problem.

Moreover, given the excess liquidity in these countries, another question that arises is whether there are (potential) borrowers in these countries that are willing to apply for a loan even at the prevailing interest rates, but are being denied?[83] Based on the credit rationing theory, rejecting loans is part of the profit maximising behaviour, because increasing credit would mean a higher default rate (Waller and Lewarne, 1994). Underlying this approach is the view that increased incidence of credit rationing, particularly of more risky borrowers, may take the form of increases in excess reserve holdings motivated, in a crisis context, by higher perceived uncertainty or risk of default. Greater volatility of deposit outflows or even large part of deposits concentrated in the few depositors may also prompt banks to keep excessive funds as a precautionary device for unforeseen contingencies (Mishkin, 2001). Therefore, the credit-rationing provides a further rationale for precautionary balances.

This situation may theoretically be explained via the backward bending loans supply curve as presented in figure 4.2 below.

[83]Credit Rationing is a situation in which there is an excess demand for commercial loans at the prevailing commercial loan rate (Jaffe and Modigliani, 1969).

Figure 4.2. Backward-bending loan supply curve



Source: Shen, 2002

r – expected return on bank lending
r* – optimal bank lending rate
L^S– loan supply curve
L^D– loan demand curve

Q_L– quantity of loans
Q_S – quantity of loans issued
Q_D – quantity of loans demanded

$Q_S = f(\overset{+}{r}, \overset{-}{default\ risk})$

As Stiglitz and Weiss (1981) argue, the lending supply curve L^S is an increasing function of the lending rate *r*, until the supply of loan reaches *r** which corresponds to the highest expected total returns. After this point, due to adverse selection, moral hazard and the resulting rising default rate, the supply of loans becomes backward-bending. The backward bending loan supply curve implies that banks select their borrowers by using the interest rate as a screening device (Waller and Lewarne, 1994). Although banks push interest rates high, since they cannot distinguish between good and bad borrowers, they refuse to lend above the interest rate *r**, even though there are still borrowers willing to pay higher rates. This is because after a certain interest rate, most low-risk borrowers that would repay the loan will not apply for loans and remaining applicants are disproportionately high-risk. Thus the borrowers that are denied credit are credit rationed and represent excess demand in the loan market.

The loan rate may be sticky because of imperfect information about potential new borrowers for reasons similar to those analyzed by Stiglitz and Weiss (1981). In particular, asymmetric information, the inability to distinguish between high and low-risk borrowers, may make banks reluctant to reduce their lending rate to attract new borrowers because of the resulting increase in the riskiness of the bank's loan portfolio. If these adverse selection effects are important enough, the loan market may not clear and banks will prefer to hold non-remunerated reserves (Saxegaard, 2006). As discussed in section 1.2, there is some limited evidence supporting these theories in transition economies. Based on an empirical study using a BEEPS dataset for 1999, 2002 and 2005, Hashi and Toçi (2010) find evidence of firms being denied credit in SEE countries. A significant proportion of small and medium sized firms in SEE countries have been facing financial constraints. From all the firms that applied for a loan, 6.5 percent were denied credit. In addition, 30.6 percent of firms did not apply for a loan, many of these were discouraged borrowers.

Not all the excess liquidity is accumulated voluntarily. For example, after the financial crisis, just like in the Euro area, in European transition countries the build-up of excess liquidity may have been a consequence of deficient borrowing due to weak growth prospects, despite relatively lower interest rates. Saxegaard (2006, p. 13) finds evidence that, amongst other factors like oil exports, the increase of excess liquidity in the sub-Saharan countries was largely induced by deficient lending or increases in government deposits "due to moral suasion to make commercial banks accept deposits even where these lead to excess liquidity". An increase in the deposit base, especially in transition countries may have been caused by capital account inflows, foreign aid, workers' remittances or even net export revenues.

Institutional factors may also encourage commercial banks to hold involuntary reserves. Weak contract enforcement and collateral execution together with a lack of an impartial judicial system, are amongst the institutional failures contributing to high information asymmetry between commercial banks and their borrowers in European transition countries (Doing Business Report, 2012). If loans are fully collaterized, the costs of defaults are diminished, however, in transition countries collateral is downgraded due to uncertain property rights and the inability to enforce contracts. These institutional problems are especially prevalent in the SEE countries. In effect, a loan will either not be issued or it will be issued at a smaller value and/or higher interest rate. These structural problems impede lending and make banks reluctant to reduce excess liquidity and extend lending.

Other studies describe the incidence of excess liquidity as a symptom of macroeconomic imbalances (Heenan, 2005). Having in mind the recent financial crisis, cyclical factors also play a crucial role in a bank's decision to lend, as recession, weak growth prospects and the need to rebuild banks' balance sheets make banks more reluctant to lend and preserve their liquid funds. For example, during the financial crisis in late 2008 and 2009, with the exception of Albania, Kosovo and Poland, all other European transition countries experienced recession, after a period of robust growth and high rates of credit growth. This way, lending may become procyclical, while liquid funds countercyclical.

To conclude, either demand or supply side driven, the key question of this research project is whether the presence of excess liquidity represents a deadweight loss for the domestic economies of the European transition economies, and if these resources were to be used in the form of lending would they contribute to more rapid growth.

The following section will discuss the measurement of excess liquidity based on the new definition, the methodology and determinants of excess liquidity in European transition economies.

4.3. Determinants of precautionary and involuntary excess liquidity

4.3.1. Model specification and methodology

In this section, the determinants of precautionary and involuntary excess liquidity in European transition countries are investigated by employing a panel framework. Our model is mainly based on the original model of Agenor et al., (2004) where banks' demand function for excess liquidity is examined and the enhanced version of Saxegaard (2006). Like Saxegaard's (2006) model, our model also distinguishes between precautionary and involuntary excess liquidity, however we use bank-level data, whereas Saxegaard used macro-level data, and furthermore, our model is augmented by additional risk and regulatory variables. The specified model will be estimated in two stages: in the first stage we employ a fixed effects model and examine the impact of various factors on excess liquidity. The parameter estimates from the first stage are used in the second stage in order to create dynamic forecasts of the levels of precautionary and involuntary excess liquidity.

In both Agenor's et al. (2004) and Saxegaard's (2006) studies, excess liquidity is measured as the level of excess reserves held by commercial banks at the central bank. We have criticised this approach for its imprecise measure of excess liquidity, because as noted in chapter 1, banks in European transition countries also hold relatively high liquidity ratios (liquid assets to total assets), even higher than the mandatory ratios However, measuring how much of the liquidity is mandatory and how much is excessive is problematic since not all central banks in European transition countries impose required liquidity ratios, Hungary only introduced them in 2011 and Albania and Kosovo have also only recently introduced them. Given that some of the countries in our dataset are well advanced in the transition process while others (SEE countries) are less so, we have taken an educated guess and arbitrarily set mandatory liquidity ratios. As discussed in chapter 1, the part of excess liquidity held above reserve required ratio (EL1) is easy to measure once data on the required reserve ratio and above mandatory reserve at central banks are available. Whereas the measurement part of the excess liquidity above mandatory liquidity ratio (EL2) is not straightforward. This is because, as discussed in section 1.4 (see footnotes 6 and 7) the mandatory liquidity ratio (e.g. liquid assets to short-term liabilities or liquid assets to total assets) was not a popular prudential measure during the 2000s, since the regulators mainly concentrated on credit risk and capital requirements. After the onset of the GFC, where the sudden drop of liquidity in banks threatened their ability to fulfil daily requirements and short-term obligations and led to many banks' bankruptcy, the importance of mandatory liquidity ratios regained central banks'attention and were re-introduced in Basel IIII. The EU member countries that implement Basel III fulfil the liquidity coverage ratio (LCR) and Net Stable Funding (NSF) ratio. Within SEE countries, Macedonia since 2011 imposed ratios of liquid assets to short-term liabilities of the same maturity of 100 percent (similar to LCR, see footnote 7); Serbia has no mandatory liquidity ratio, whereas Kosovo and Albania impose a liquid assets to short-term liabilities ratio of 25 percent. The rest of the countries in our sample still have no mandatory liquidity ratios. Instead of taking the typical measure of liquid assets to short-term liabilities indicator, and in order to be able to sum up EL1 and EL2, we needed to come up with a measure with the same denominator, which is 'total assets'. For the SEE countries, given that short-term liabilities represent a smaller volume than total banking assets, we decided that our mandatory liquidity ratio, i.e. liquid assets to total assets should be smaller than 25 percent and was set to 20 percent. Other more advanced transition countries were considered to be less risky and have more developed banking system and hence a lower mandatory ratio of 15 percent was chosen. From here, the part that banks hold above than 20 percent (15 percent), was treated as excess liquid assets (e.g. liquid assets/total assets 20 percent = EL2). This part of excess reserves was added to the excess liquidity held at central banks (e.g. EL1 + EL2). We have calculated the excess liquidity at central banks as a ratio to total assets, in order to have the same denominator so that both parts of excess liquidity can be added together:

$$Total\ EL = \frac{EL\ at\ Central\ Banks}{Total\ Assets} + \frac{EL\ above\ Mandatory\ Liquidity\ Ratio}{Total\ Assets} = EL1 + EL2 \qquad (4.16)$$

Therefore, the dependent variable in our research project includes both parts of excess liquidity. As in Saxegaard (2006), the precautionary and involuntary excess liquidity will be specified separately.

The description of the rest of the explanatory variables in precautionary and involuntary excess liquidity (equation 4.16 and 4.17 below) is explained in table 5.1 of section 5.4. The precautionary and involuntary excess liquidity models are specified as in the following:

$$EL^{p}_{it} = \alpha + \alpha_1 RR_{i,t} + \alpha_2 i^{D}_{i,t} + \alpha_3 Eq_{i,t} + \alpha_4 gdp1_{i,t} + \alpha_5 Volc_{i,t} + \alpha_6 Vold_{i,t} + \alpha_7 Vold_{i,t} + \alpha_8 npl_{i,t} + \varepsilon_{i,t} \qquad (4.17)$$

In equation (4.15) EL^{p}_{it} denotes the precautionary excess liquidity, α is the intercept, $RR_{i,t}$ denotes the required reserves ratio set by the central bank, $i^{D}_{i,t}$ is the interest rate on deposits, $Eq_{i,t}$ represents the total bank equity to total assets, $gdp1_{i,t}$ denotes the $t + 1$ growth rate GDP of the respective country, $VolC_{i,t}$, $VolD_{i,t}$, and $VolG_{i,t}$ represent the five year moving averages of the standard deviation of the private sector credit to GDP, deposits to GDP and the volatility of GDP, respectively, and $npl_{i,t}$ denote the non-performing loans to total loans ratio for each bank; i denotes each bank, t denotes each year and ε is the error term.

The required reserves ratio set by the central bank, represents an explanatory variable in the precautionary equation, whose effect may go either way, depending on the circumstances. Saxegaard (2006) argues that an increase in the required reserve ratio, particularly in the countries that gain no remuneration on required reserves, would increase the cost of holding other excess reserves, thus induce them to hold less excess liquidity. In contrast, Aikaeli (2006) argues that when required reserves tend to change more often, banks respond by holding more excess reserves to anticipate changes from the regulator. On the other side of the precautionary variables, we include the deposit interest rate as a proxy for the cost of liquidity for banks. Agenor et al. (2004) and Saxegaard (2006) use the discount window or the market rate as a proxy for liquidity cost. However, not all the countries in our sample have developed capital markets and in some other countries (e.g. Kosovo and Montenegro) monetary policy lacks such instruments. Hence, the liquidity cost is proxied by the deposit interest rate the banks's costs of holding customer deposits. Furthermore, given that the interbank market in transition economies is typically underdeveloped, banks have to borrow from depositors as it represents their main funding source. Thus, in this respect our model has been modified to adapt to the transition context. The theory suggests that, other things being equal, banks to hold a larger amount of excess reserves if the cost of borrowing at the discount window is high.

Another potentially important factor that was neglected in other studies is the equity to total assets ratio. As equity represents a cushion against malfunction, this ratio measures the protection afforded to the bank by the equity banks invested in (Bankscope, 2013).[84]

The higher this ratio is, the more protection banks have to enlarge investments, thus the less need for excess liquidity.[85] Next year's GDP growth rate is used as a proxy for the banks' expectation towards economic prospects of the country. An expectation of positive growth in the upcoming year is expected to influence banks' decision on lending more in the current year, and thus hold less excess liquidity; whereas negative expectations are expected to have the opposite effect. Theory suggests that the $Volc_{i,t}$, $Vold_{i,t}$, and $Vold_{i,t}$ act as risk measures; thus, an increase in the volatility of these indicators, other things being equal, would motivate banks to increase their demand for precautionary excess reserves. Another typical risk measure included in our model is the non-performing loans to total loans ratio. This may account for both, the riskiness of the economic environment, which is the ability of clients to repay loans, and the level of bank's managerial skills to monitor and screen loan applicants (Podpiera and Weill, 2008; IMF, 2012). An increase in the non-performing loans ratio, other

[84] Equity in levels (or often referred to as capital) represents the difference between assets and liabilities; it usually includes share capital, profits and reserves. Based on Bankscope's definition and calculation, equity includes common shares and premium; retained earnings; reserves for general banking risks and statutory reserves. A bank's capital can be thought of as the margin to which creditors are covered if a bank liquidates its assets.

[85] In chapter 3, the output gap was estimated with a Kalman filter for Kosovo, Czech Republic and Estonia. The explicit relationship between output gap and excess liquidity will be examined later in this chapter, in another model utilizing macro-level data. In chapter 5 we find that the relationship between output gap and excess liquidity (or as in Saxegaard's model the relationship between excess liquidity and the output gap), is not actually causal. Therefore, adding one as a determinant of the other may actually represent a misspecification of the equation.

things being equal, is expected to motivate banks to hold more excess liquidity as a precaution against further loss.

$$EL^{I}_{it} = \beta + \beta_1 Dep_{i,t} + \beta_2 i^{L}_{i,t} + \beta_3 Cred_{i,t} + \beta_4 Leg_{i,t} + \beta_5 Fdi_{i,t} + \beta_6 i^{2L}_{i\,i,t} + v_{i,t} \qquad (4.18)$$[86]

In equation (4.16) EL^{I}_{it} denotes the involuntary excess liquidity, β is the intercept, $Dep_{i,t}$ represents customer deposits to GDP, $i^{L}_{i,t}$ denotes the loan interest rate, whereas i^{2L} denotes the squared lending interest rate; $Cred_{i,t}$ represents private sector credit to GDP, $Leg_{i,t}$ is a macro-level index measuring the strength of legal rights, ranging from 0 to 10 (0 = weak to 12 = strong legal rights); the higher the score the better the laws in that country to facilitate lending. $Fdi_{i,t}$, is also macro-level factor, denoting the foreign direct investment to GDP.

Except for the Saxegaard (2006) study, there is a lack of theoretical guidance on the factors that explain the size of involuntary excess liquidity. Deposits at commercial banks are usually the main source of excess liquidity. This may be the case particularly in less developed transition countries, where banks conduct a simple banking model with domestic deposits representing the main funding source: as the level of deposits increase, excess liquidity also increases. Gilmour (2005) explains that excess liquidity in the banking system during transition by the need for businesses to maintain large liquid balances for operational as well as investment needs, given the difficulty of obtaining credit. The banks' lending rate is expected to be positively correlated with excess liquidity, especially in the European transition countries (see chapter 1) where loan interest rates are already quite high. An increase in the lending rate would, other things being equal, reduce lending and increase excess liquidity (see the backward-bending supply curve in the previous section). Agenor et al. (2004) and Wyplosz (2005) argue that weak bank lending, either due to credit rationing or lack of absorption capacity, may also be one of the causes behind the accumulation of excess liquidity. One of the most important and a likely determinant to consider with regard to our sample of countries are institutional factors. As argued earlier in this chapter, institutional factors may affect the information asymmetry between banks and lenders. The effectiveness of these factors, such as the laws designed to facilitate business, is outside commercial banks' control, thus they are exogenous in nature. These factors are proxied by the World Bank's

[86] In the involuntary excess liquidity equations, we have also controlled for the effect of remittances to GDP (macro-level variable sourcing from World Bank database). However, this variable turned out as statistically insignificant, the model did not perform well and it was highly correlated with deposits to GDP variable, thus it was dropped from equation.

Doing Business (2013) index, i.e. Strength of Legal Rights Index (LEG), which measures the degree to which collateral and bankruptcy laws protect the rights of borrowers and lenders and thus facilitate lending. The index ranges from 0 to 12, with higher scores indicating that these laws are better designed to expand access to credit, however the Index does incorporate the extent to which the legislation is enforced.[87] Finally, the remaining explanatory variable is foreign direct investment inflows (FDI). This is because FDIs are large in many European transition countries (e.g. Estonia, Hungary and Bulgaria receive on average annually net inflows of FDIs equivalent to 10 per cent of their GDP, whereas in Montenegro they are up to 23% GDP), so they may increase the deposits base in banks. Finally, it should be noted that the theory on excess liquidity determinants does not proclaim a non-linear relationship between excess liquidity and lending interest rate (Agenor et al., 2004; Saxegaard, 2006; Khemrraj, 2007 etc.). However, the credit rationing theory proclaims that after certain point in lending interest rate, the loan supply curve becomes backward-bending, meaning that the relationship is no longer linear. Having said this, we will try both combinations of involuntary excess liquidity equations, with linear and non-linear relationship between excess liquidity and lending interest rate, and based on the result decide which one applies in the case of our sample.

Due to data limitations, some of other important variables considered in the theory could not be investigated, such as the cash to deposit ratio to account for liquidity risk. In addition, following Saxegaard (2006) a very important variable to employ in our study was thought to be the volatility of the output gap for each country; however, as was noted in the previous chapter, a measure of the output gap is not available for all the countries in our sample

The precautionary and involuntary models utilise bank-level data from 84 European transition countries banks, sourcing from the international database, Bankscope. In the involuntary excess liquidity the *Leg* and *Fdi* variables factors are taken from the World Bank database.. However, due to the relatively high rate of the missing values in the database, we are left with around 100 observations for each model. It should be noted that the missing values appear to be randomly allocated, they do not follow a pattern or relate to a particular variable, but are spread throughout the database. The rate of missing values is higher in the earlier years of the

[87] Other studies that examine the causes of excess liquidity acknowledge the importance of institutional factors, but admit that they were impossible to control for due to the lack of data. This was the case with Saxegaard's study published in 2006. The time series on the strength of legal rights index starts only from 2007, thus it is relatively short.

time series (1998-2012), which is to be expected for the countries in our sample, as some of them started the compilation of official statistics later in the transition process (in line with the process of the institutional building process), and furthermore the quality of the data is expected to be better for the later years.

One important issue that should be discussed before estimating the model is potential endogeneity due to missing variables.[88] An important variable that could be added into our model, is the quality of managerial decisions in the banks. With this variable we could control the impact of decision making in the banks on how much to lend and how much to hold as excess reserves. However, a proxy for this factor is to the best of our knowledge is not available. This factor may be correlated with the dependent variable, the loans issued and presumably the non-performing loans ratio.[89] The problem with an omitted variable in the model is that it biases the estimates and generates correlation of error term with explanatory variables. However, as we will see in what follows, the endogeneity problem in our model will be addressed via adopting the fixed effects method.

In other studies, possible endogeneity arising from the simultaneity problem was not addressed. However, based on the practitioner's knowledge, we will argue that our model does not suffer from the simultaneity problem, or in the cases where simultaneity may be an issue, it is of negligible nature. Starting from the precautionary model, $RR_{i,t}$, $Volc_{i,t}$, $Vold_{i,t}$ and $\alpha_7 Volg_{i,t}$ are treated as exogenous factors. The deposit interest rate may be an important factor causing the excess liquidity to change, however the excess liquidity rate may not have a direct link with the deposit interest rate. This is because, the deposit interest rate, first of all is dependent on other factors that have nothing to do with excess liquidity, such as administrative and personnel costs, competitiveness of the banking market, creating a long-term relationship with their customers etc., meaning a direct simultaneous relationship between excess liquidity and deposit interest rates is unlikely. The GDP of the next year may affect the banks' decision on their annual planning, however excess liquidity ratio is not expected to impact the level of the next year's GDP. The equity to total assets is also not expected to be affected by the level of excess liquidity. This is because equity comprises the

[88] The omitted variable problem is when there is one or more variables that we would like to hold fixed when estimating the ceteris paribus effect on explanatory variable (Wooldrigde, p.490).
[89] Podpiera and Weill (2008) define non-performing loans as a conventional risk taking measure which proxies the quality of the operating environment, whereas Berger and DeYoung (1997) also point out that poor managerial skills and improper loan monitoring leads to inefficient loan allocation.

initial capital invested by the shareholders, which is partially based on regulatory requirements, thus excess liquidity cannot cause changes in the ratio of equity to total assets. The non-performing loans may be an important factor causing the build-up of excess reserves, especially in a crisis time. However, non-performing loans to a large extent depend on the business cycle phase of the economy, inability to make loan repayments by customers, the managerial skills of the bank to monitor and screen the loan applicant, hence a simultaneous relationship between excess liquidity and non-performing loans is unlikely.

With regard to the involuntary model, the strength of the legal rights, FDIs represent macroeconomic factors that are outside of the individual banks control, thus are assumed as exogenous. The loan interest rate may cause a build-up of excess liquidity, especially in the cases when loan interest rates are already high as is the case in many transition countries. However, loan interest rates depend on risk factors such as inflation, bad economic environment, exchange rate uncertainty etc., financial costs, operational costs and tariffs and commissions that have nothing to do with the level of excess liquidity. The deposits to GDP ratio cause the build-up of excess liquidity accumulation, however in the balance sheet of banks, deposits are proportionally much larger than excess liquidity accumulated. Further, the accumulation of deposits depends on the decision of depositors: when to deposits, how much to deposit and with what maturity date. Thus, excess liquidity cannot cause the accumulation of deposits. The credit to GDP ratio, as in the case of a credit crunch may cause banks to hold more excess liquidity, due to higher uncertainties in loan repayment.

In the following, the methodology for estimating determinants of precautionary and involuntary excess liquidity will be presented.

The bank level excess liquidity model is estimated via the fixed effects (FE) technique, which is generally a more appropriate approach when the dataset is small, as in our case. The reason for using fixed effects rather than a random effects model is also based on the variety of the countries in our sample, starting from the less developed transition countries such as Kosovo, Montenegro and Bosnia up to the most developed European transition countries, which are already part of the European Union (e.g. the Czech Republic, Estonia, Hungary etc.) This variety will be captured in the specific country intercepts enabled in the fixed effects method. In addition, the detailed analysis of the data below, including possible multicollinearity and endogeneity problems, suggests that the FE model should be employed for the econometric analysis, whereas a problem with the random effects is that it precludes the inclusion of a lagged dependent variable, which may be required if the model suffers from autocorrelation.

A major attraction of the fixed effects method is the ability to control for all stable characteristics of the banks, thereby eliminating potentially large sources of bias. As with first differencing of the equation, fixed effects use a transformation to remove the unobserved time-invariant part of each bank specific α_i (in our case α_i for precautionary and β_i for involuntary motive) prior to estimation. When the unobserved effects are correlated with the explanatory variables the fixed effects method helps to avoid omitted variable bias due to unobserved heterogeneity, by taking the unobserved part in the excess liquidity model from the error term into the observable part of the model (Studenmund, 2006). This is because, by differencing out the individual variability across banks, the fixed effects methods eliminates much of the error variance that is present and hence produces a more powerful test hypothesis. The rationale is that by differencing out the individual variablilty across banks, FE eliminates much of the error variance that is present (Buddelmeyer et al., 2008). Nevertheless, when using fixed effects there is a trade-off between the advantage of eliminating bias in the estimates to losing the between banks variability. The fixed effects approach has the disadvantage of not being able to estimate the coefficients of the stable variables that have no within-banks variance, as all the within-banks invariant indicators are eliminated. However, fixed effects controls for stable variables even when there are no measures available for them.

In the fixed effects estimation time dummy variables are included, that is yearly variables from 1999-2012, to allow for aggregate time effects. The time dummy variables are treated as exogenous, 'because the passage of the time is exogenous – and so they act as their own instruments' (Wooldridge, p.490).

The diagnostic test from the static fixed effects specification, – rho,- suggests that the model suffers from residual autocorrelation. The presence of error autocorrelation may indicate that other important unobserved dynamic factors that explain excess liquidity are omitted from our model. Before assuming that a dynamic panel specification should be pursued (differenced or GMM), it may be possible that in a small model the dynamics are unobservable and contained within residual. In an attempt to address the residual autocorrelation that is by modelling the dynamics in the error terms we pursued the Cochrane-Orcutt method, which is presented in the following section.

4.3.2. Deriving and testing the common factor restrictions

Before continuing with the estimations, some prior conditions for model validity must be met: the so called *common factor restrictions* (CFR) must hold. In order to test for CFRs, the model specification should be transformed into a dynamic linear regression model. The general approach to deriving the CFRs is presented in box 4.1.

Box 4.1. Common Factor Restrictions Model

In order to test for CFRs, it is assumed that in the static model, at least one explanatory variable has an autoregressive error term of order one – AR(1):

$Y_{i,t} = \alpha + \alpha_1 X_{it} + \varepsilon_{it}$ (4.19)

where $\varepsilon_{i,t} = \rho\varepsilon_{i,t-1} + v_{i,t}$ and $0 < \rho < 1$ (4.20)

In order to tackle the autocorrelation of the error term, we add dynamics into the linear regression model (4.17) following the Cochrane-Orcutt approach:

$Y_{i,t} = (1-\rho)\alpha + \rho Y_{i,t-1} + \alpha_1 X_{i,t} - \rho X_{i,t-1} + v_{i,t}$ (4.21)

Equation (4.21) represents a restricted version of the following dynamic linear. Here, the slope coefficients of a static model are estimated conditional on an AR(1) dynamic linear regression model (Sargan, 1964):

$Y_{i,t} = \alpha + \alpha_1 Y_{i,t-1} + \alpha_2 X_{i,t} + \alpha_3 X_{i,t-1} + v_{i,t}$ (4.22)

The dynamic linear regression in (4.22) can be transformed into its restricted version (4.21), if and only if the common factor restrictions hold in the equation (4.22), that is:

$\alpha_1 * \alpha_2 = - \alpha_3$

That is, the slope coefficient of each continuous variable in the static model are estimated conditional on the AR(1) dynamic in the residual. If the CFRs hold, then an unobserved components model should be specified, because under this assumption a pure serial correlation in the residuals can be assumed.

Based on the derivation in box 4.1., the CFRs will next be derived step-by-step and tested with respect to our specific excess liquidity model.

a) CFRs for Precautionary Motive

$EL^{p}_{it} = \alpha + \alpha_1 RR_{i,t} + \alpha_2 i^{D}_{i,t} + \alpha_3 Eq_{i,t} + \alpha_4 gdp1_{i,t} + \alpha_5 Volc_{i,t} + \alpha_6 Vold_{i,t} + \alpha_7 Vold_{i,t} + \alpha_8 npl_{i,t} + \varepsilon_{i,t}$ (4.23)

Where $\varepsilon_{i,t} = \rho\varepsilon_{i,t-1} + v_{i,t}$ (4.24)

1) Transformation into dynamic linear regression (lag once):

$$EL^{p}_{i,t-1} = \alpha+\alpha_1 RR_{i,t-1}+ \alpha_2 i^{D}_{i,t-1}+ \alpha_3 Eq_{i,t-1}+ \alpha_4 gdp1_{i,t-1}+ \alpha\ _5 Vol_{i,t-1}+ \alpha_6 Vold_{i,t-1}+ \alpha_7 Volg_{i,t-1}+ \alpha_8 npl_{i,t-1}+\varepsilon_{i,t-1} \quad (4.25)$$

2) Solve for $\varepsilon_{i,t-1}$

$$\varepsilon_{i,t-1} = EL_{i,t-1}- \alpha -\alpha_1 RR_{i,t-1}- \alpha_2 i^{D}_{i,t-1}- \alpha_3 Eq_{i,t-1}- \alpha_4 gdp1_{i,t-1}- \alpha\ _5 Volc_{i,t-1}- \alpha_6 Vold_{i,t-1}- \alpha_7 Volg_{i,t-1}- \alpha_8 npl_{i,t-1} \quad (4.26)$$

3) Substitute (4.26) into (4.24)

$$\varepsilon_{i,t} =\rho\ (EL^{p}_{i,t-1}- \alpha -\alpha_1 RR_{i,t-1}- \alpha_2 i^{D}_{i,t-1}- \alpha_3 Eq_{i,t-1}- \alpha_4 gdp1_{i,t-1}- \alpha\ _5 Volc_{i,t-1}- \alpha_6 Vold_{i,t-1}- \alpha_7 Volg_{i,t-1}- \alpha_8 npl_{i,t-1})+ v_{i,t} \quad (4.27)$$

$$\varepsilon_{i,t} =\rho ELV_{i,t-1}- \rho\alpha- \rho\alpha_1 RR_{i,t-1}- \rho\alpha_2 i^{D}_{i,t-1}- \rho\alpha_3 Eq_{i,t-1}- \rho\alpha_4 gdp1_{i,t-1}- \rho\alpha\ _5 Volc_{i,t-1}- \rho\alpha_6 Vold_{i,t-1}- \rho\alpha_7 Volg_{i,t-1}- \rho\alpha_8 npl_{i,t-1} +v_{i,t} \quad (4.28)$$

4) Substitute (4.26) into (4.21)

$$EL^{p}_{it} = \alpha+\alpha_1 RR_{i,t}+ \alpha_2 i^{D}_{i,t}+ \alpha_3 Eq_{i,t}+ \alpha_4 gdp1_{i,t}+ \alpha\ _5 Volc_{i,t}+ \alpha_6 Vold_{i,t}+ \alpha_7 Volg_{i,t}+ \alpha_8 npl_{i,t}+ \rho\ EL_{i,t-1}- \rho\ \alpha- \rho\ \alpha_1 RR_{i,t-1}- \rho\ \alpha_2 i^{D}_{i,t-1}- \rho\ \alpha_3 Eq_{i,t-1}- \rho\ \alpha_4 gdp1_{i,t-1}- \rho\alpha_5 VOLc_{i,t-1}- \rho\ \alpha_6 Vold_{i,t-1}- \rho\ \alpha_7 Volg_{i,t-1}- \rho\ \alpha_8 npl_{i,t-1} +v_{i,t} \quad (4.29)$$

5) Collect terms, hence:

$$EL^{p}_{it}=(1- \rho)\alpha+ \alpha_1 RR_{i,t}+ \alpha_2 i^{D}_{i,t}+ \alpha_3 Eq_{i,t}+ \alpha_4 gdp1_{i,t}+ \alpha\ _5 Volc_{i,t}+ \alpha_6 Vold_{i,t}+ \alpha_7 Volg_{i,t}+ \alpha_8 npl_{i,t}+ \rho EL_{i,t-1} - \rho\alpha_1 RR_{i,t-1}- \rho\alpha_2 i^{D}_{i,t-1} - \rho\alpha_3 Eq_{i,t-1} - \rho\alpha_4 gdp1_{i,t-1}- \rho\alpha\ _5 Volc_{i,t-1}- \rho\alpha_6 Vold_{i,t-1}- \rho\alpha_7 Volg_{i,t-1}- \rho\alpha_8 npl_{i,t-1} +v_{i,t} \quad (4.30)$$

Ignoring the constant term α, equation (4.30) has nine coefficients: $\rho,\ \alpha_1, \alpha_2, ,\alpha_3, \alpha_4, \alpha_5,\ \alpha_{6,}\alpha_7\ \alpha_8$ Equation (4.30) is a restricted version of a dynamic linear regression model equation (4.31)-unrestricted, which has 17 coefficients: $\rho,\ \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8, \alpha_9, \alpha_{10}, \alpha_{11}, \alpha_{12}, \alpha_{13}, \alpha_{14}, \alpha_{15},\ \alpha_{16}.$

$$EL^{p}_{it}= \alpha+ \alpha_1\ EL_{i,t-1} + \alpha_2 RR_{i,t}+ \alpha_3 i^{D}_{i,t}+ \alpha_4 Eq_{i,t}+ \alpha_5 gdp1_{i,t}+ \alpha_6 Volc_{i,t}+ \alpha_7 Vold_{i,t}+ \alpha_8 Volg_{i,t}+ \alpha_9 npl_{i,t}+ \alpha_{10} RR_{i,t-1}+ \alpha_{11} i_{i,t-1}+ \alpha_{12} Eq_{i,t-1}+ \alpha_{13} gdp1_{i,t-1}+ \alpha\ _{14} Volc_{i,t-1}+ \alpha_{15} Vold_{i,t-1}+ \alpha_{16} Volg_{i,t-1}+ \alpha_{17} npl_{i,t-1}+ \varepsilon_{i,t} \quad (4.31)$$

→ In both (4.30) and (4.31) there is one coefficient on $EL_{i,t-1}$, which is ρ from (4.30) and α_1 from (4.31)

→ in (4.30) the coefficient on $RR_{i,t-1}$ is $-\rho\alpha_1$, the coefficient on $i_{i,t-1}$ is $-\rho\alpha_2$, *on* $Eq_{i,t-1}$ is $-\rho\alpha_3$, on $gdp1_{i,t-1}$ is $-\rho\alpha_4$, in $VolC_{i,t-1}$ is $-\rho\alpha_5$, in $Vold_{i,t-1}$ is $-\rho\alpha_6$, in $Volg_{i,t-1}$ is $-\rho\alpha_7$, in $npl_{i,t-1}$ is $-\rho\alpha_8$

→ in (4.31) the coefficient on $RR_{i,t-1}$ is α_9, the coefficient on $i_{i,t-1}$ is α_{10}, on $Eq_{i,t-1}$ is α_{11}, on $gdp1_{i,t-1}$ is α_{12}, $Volc_{i,t-1}$ is α_{13}, in $Vold_{i,t-1}$ is α_{14}, in $Volg_{i,t-1}$ is α_{15}, in $npl_{i,t-1}$ is α_{16}

Hence, the $-\rho\alpha_1$ is the negative product of the coefficients on $EL_{i,t-1}$ and $RR_{i,t-1}$, and $-\rho\alpha_2$ is the negative product of the coefficients on $EL_{i,t-1}$ and $i_{i,t-1}$, $-\rho\alpha_3$ is the negative product of $EL_{i,t-1}$ and $Eq_{i,t-1}$ and so on…

H^P_o: the dynamic linear regression model (4.31) can be transformed into model (4.30), if and only if in the dynamic linear regression model (4.31) the following restrictions hold:

The null hypothesis of common factor restrictions for each coefficient (single tests) and the whole model (multiple test) is defined in the following way:

$-\alpha_{10} = \alpha_1 * \alpha_2$ single test for *RR*

$-\alpha_{11} = \alpha_1 * \alpha_3$ single test for i^D

$-\alpha_{12} = \alpha_1 * \alpha_4$ single test for *Eq*

$-\alpha_{13} = \alpha_1 * \alpha_5$ single test for *gdp1*

$-\alpha_{14} = \alpha_1 * \alpha_6$ single test for *Volc*

$-\alpha_{15} = \alpha_1 * \alpha_7$ single test for *Vold*

$-\alpha_{16} = \alpha_1 * \alpha_8$ single test for *Volg*

$-\alpha_{17} = \alpha_1 * \alpha_9$ single test for *npl*

these are *common factor restrictions (CFR)* for Precautionary Motive Model

$(-\alpha_{10} = \alpha_1 * \alpha_2)(-\alpha_{11} = \alpha_1 * \alpha_3)(-\alpha_{12} = \alpha_1 * \alpha_4)(-\alpha_{13} = \alpha_1 * \alpha_5)(-\alpha_{14} = \alpha_1 * \alpha_6)(-\alpha_{15} = \alpha_1 * \alpha_7)(-\alpha_{16} = \alpha_1 * \alpha_8) -\alpha_{17} = \alpha_1 * \alpha_9)$, multiple tests for the whole model.

b) CFRs for Involuntary Motive

$$EL^I_{it} = \beta + \beta_1 Dep_{i,t} + \beta_2 i^L_{i,t} + \beta_3 Cred_{i,t} + \beta_4 Leg_{i,t} + \beta_5 Fdi_{i,t} + \beta_{6i}\, i^{2L}_{i,t,t} + u_{i,t} \quad (4.32)$$

Where $u_{i,t} = \rho u_{i,t-1} + v_{i,t}$ (4.33)

1) Transformation into a dynamic linear regression (Lag once):

$$EL^I_{i,t-1} = \beta + \beta_1 Dep_{i,t-1} + \beta_2 i^L_{i,t-1} + \beta_3 Cred_{i,t-1} + \beta_4 Leg_{i,t-1} + \beta_5 Fdi_{i,t-1} + \beta_6 i^{2L}_{i\ \ i,t-1} + u_{i,t-1} \quad (4.34)$$

2) **Solve for $u_{i,t-1}$:**

$$u_{i,t-1} = EL^{I}_{i,t-1} - \beta - \beta_1 Dep_{i,t-1} - \beta_2 i^{L}_{i,t-1} - \beta_3 Cred_{i,t-1} - \beta_4 Leg_{i,t-1} - \beta_5 Fdi_{i,t-1} - \beta_6 i^{2L}_{i\ i,t-1} \quad (4.35)$$

3) **Substitute (4.35) into (4.33):**

$$u_{i,t} = \rho(EL^{I}_{i,t-1} - \beta - \beta_1 Dep_{i,t-1} - \beta_2 i^{L}_{i,t-1} - \beta_3 Cred_{i,t-1} - \beta_4 Leg_{i,t-1} - \beta_5 Fdi_{i,t-1} - \beta_6\, i^{2L}_{i\ i,t-1}) + v_{i,t} \quad (4.36)$$

$$u_{i,t} = \rho EL^{I}_{i,t-1} - \rho\beta - \rho\beta_1 Dep_{i,t-1} - \rho\beta_2 i^{L}_{i,t-1} - \rho\beta_3 Cred_{i,t-1} - \rho\beta_4 Leg_{i,t-1} - \rho\beta_5 Fdi_{i,t-1} - \rho\beta_6\, i^{2L}_{i,t-1} + v_{i,t} \quad (4.37)$$

4) **Substitute (4.37) into (4.32)**

$$EL^{I}_{it} = \beta + \beta_1 Dep_{i,t} + \beta_2 i^{L}_{i,t} + \beta_3 Cred_{i,t} + \beta_4 Leg_{i,t} + \beta_5 Fdi_{i,t} + \beta_6\, i^{2L}_{i,t} + \rho EL^{I}_{i,t-1} - \rho\beta - \rho\beta_1 Dep_{i,t-1} - \rho\beta_2 i^{L}_{i,t-1} - \rho\beta_3 Cred_{i,t-1} - \rho\beta_4 Leg_{i,t-1} - \rho\beta_5 Fdi_{i,t-1} - \rho\beta_6 i^{2L}_{i,t-1} + v_{i,t} \quad (4.38)$$

5) **Collect terms, hence:**

$$EL^{I}_{it} = (1-\rho)\beta + \beta_1 Dep_{i,t} + \beta_2 i^{L}_{i,t} + \beta_3 Cred_{i,t} + \beta_4 Leg_{i,t} + \beta_5 FDdi_{,t} + \beta_6 i^{2L}_{i,t} + \rho EL^{I}_{i,t-1} - \rho\beta_1 Dep_{i,t-1} - \rho\beta_2 i^{L}_{i,t-1} - \rho\beta_3 Cred_{i,t-1} - \rho\beta_4 Leg_{i,t-1} - \rho\beta_5 Fdi_{i,t-1} - \rho\beta_6 i^{2L}_{i,t-1} + v_{i,t} \quad (4.39)$$

Ignoring constant term β, equation (4.39) has seven coefficients: $\rho, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$

Equation (4.39) is a restricted version of a dynamic linear regression model (4.40) – unrestricted –, which has fourteen coefficients: $\rho, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta, \beta_8, \beta_9, \beta_{10}, \beta_{11}, \beta_{12}, \beta_{13}$

$$EL^{I}_{it} = \beta + \beta_1 EL^{I}_{i,t-1} + \beta_2 DEP_{i,t} + \beta_3 i^{L}_{i,t} + \beta_4 CRED_{i,t} + \beta_5 LEG_{i,t} + \beta_6 FDI_{i,t} + \beta_7 i^{2L}_{i,t} + \beta_8 DEP_{i,t-1} + \beta_9 i^{L}_{i,t-1} + \beta_{10} CRED_{i,t-1} - \beta_{11} LEG_{i,t-1} - \beta_{12} FDI_{i,t-1} - \beta_{13} i^{2L}_{i,t-1} + u_{i,t} \quad (4.40)$$

→ In both (17) and (18) there is one coefficient on $EL_{i,t-1}$, which is ρ from (4.39) and β_1 from (4.40)

→ in (4.37) the coefficient on $DEP_{i,t-1}$ is $-\rho\beta_1$, the coefficient on $i^{L}_{i,t-1}$ is $-\rho\beta_2$, on $Cred_{i,t-1}$ is $-\rho\beta_3$, on $\rho\beta_4 Leg_{i,t-1}$ is $-\rho\beta_4$, in $Fdi_{i,t-1}$ is $-\rho\beta_5$ and in $i^{2L}_{i,t-1}$ is $-\rho\beta_6$

→ in (4.38) the coefficient on $Dep_{i,t-1}$ is β_7, the coefficient on $i^{L}_{i,t-1}$ is β_8, on $Cred_{i,t-1}$ is β_9, on $Leg_{i,t-1}$ is β_{10}, on $Fdi_{i,t-1}$ is β_{11} and on $i^{2L}_{i,t-1}$ is β_{12}

Hence, the $-\rho\beta_1$ is the negative product of the coefficients on $EL^I_{i,t-1}$ and $Dep_{i,t-1}$, and $-\rho\beta_2$ is the negative product of the coefficients on $EL^I_{i,t-1}$ and $i^L_{i,t-1}$, $-\rho\beta_3$ is the negative product of $EL^I_{i,t-1}$ and $Cred_{i,t-1}$ and so on…..

H^I_o: The dynamic linear regression model (4.40) can be transformed into (4.39), if and only if in the dynamic linear regression model (4.40) the following restrictions hold:

The null hypothesis of common factor restrictions for each coefficient (single tests) and the whole model (multiple test) is defined in the following way:

$$\left.\begin{array}{l} -\beta_8 = \beta_1 * \beta_2 \text{ single test for } Dep \\ -\beta_9 = \beta_1 * \beta_3 \text{ single test for } i^L \\ -\beta_{10} = \beta_1 * \beta_4 \text{ single test for } Cred \\ -\beta_{11} = \beta_1 * \beta_5 \text{ single test for } Leg \\ -\beta_{12} = \beta_1 * \beta_6 \text{ single test for } Fdi \\ -\beta_{13} = \beta_1 * \beta_7 \text{ single test for } i^{2L} \end{array}\right\} \begin{array}{l} \text{these are } \textit{common factor restrictions (CFR)} \\ \text{for Involuntary Motive Model} \end{array}$$

$(-\beta_8 = \beta_1 * \beta_2)(-\beta_9 = \beta_1 * \beta_3)(-\beta_{10} = \beta_1 * \beta_4)\ (-\beta_{11} = \beta_1 * \beta_5)(-\beta_{12} = \beta_1 * \beta_6)(-\beta_{13} = \beta_1 * \beta_7)$ multiple tests for the whole model.

The dynamic model is misspecified in the panel context. Accordingly, two approaches were used to estimate the first order dynamic panel model and then tested for CFRs: OLS and FE. In OLS the coefficient on the lagged dependent variable is subject to maximum *upward* bias and in the fixed effects estimation the coefficient on the lagged Dependent variable is a subject to maximum *downward* bias. Both models are tested for CFRs for the purpose of accounting for the whole range of possible dynamic misspecification (Bond, 2002).

The tests for common factor restrictions are presented in table 4.1 (see Appendix 4.1 for the full output of results). The results for precautionary excess liquidity are presented in panel a). From the pooled OLS estimation of the dynamic linear regression of order one, the null hypothesis cannot be rejected in any of the single tests for each dependent variable, except for the equity to capital *Eq* variable which most probably affects the joint test. In a similar fashion, CFRs hold for each explanatory variable in FE specification, however, the joint test of all in dependent variables is rejected, most probably due to the volatility of deposits – *Vold*

- (CFR hold at 1% and 5% but not at 10% level of significance). In panel b) where CFRs tests for involuntary excess liquidity are presented, the results are less consistent. In the pooled OLS, the CFRs hold for all variables, except for the strength of legal rights variable – *Leg* – in both specifications (lending interest rate specified as linear and non-linear). Similarly, the CFR for this variable do not hold in FE model, in addition to the Foreign Direct Investment to GDP – *Fdi* - CFRs hold. Conclusively, given that most of the CFR in single tests hold, the static fixed effects model conditional on AR(1) error term specification is the correct model specification for estimating the determinants of excess liquidity.

Table 4.1. Tests for common factor restrictions of the dynamic linear regression, p-values of Wald test

Model	Single tests								Multiple tests
a) Precautionary	***RR***	i^D	***Eq***	***gdp1***	***VolC***	***VolD***	***VolG***	***npl***	
Pooled OLS	0.743	0.359	**0.042**	0.125	0.924	0.073	0.845	0.118	0.000
Fixed Effects	0.937	0.525	0.133	0.227	0.798	0.077	0.322	0.506	0.000
b) Involuntary	***DEP***	i^L	***CRED***	***LEG***	***FDI***	i^{2L}			
Pooled OLS	0.119	0.127	0.6	**0.013**	0.204				0.000
Pooled OLS, iloans2	0.132	0.713	0.670	**0.017**	0.257	0.304			0.000
Fixed Effects	0.697	0.254	0.510	**0.023**	0.087				0.002
Fixed Effects, iloans2	0.546	0.631	0.419	**0.017**	**0.044**	0.507			0.000

4.3.3. Discussion of the results of the model

Prior to the discussion of the models' findings, the diagnostics of the precautionary and involuntary models are checked. The initial diagnostic test we are interested in is related to the estimated coefficient of ρ. In other words, the null hypothesis tested is that the residuals are serially in dependent, against the alternative that the residuals are generated by a stationary first order autoregression (Bhargava et al., 1982):

$|\rho| < 0$, that is:

H_o: residuals are serially independent

H_1: residuals are generated by a first-order autoregression - AR(1)

The Baltagi-Wu LBI test is the best invariant test for serial correlation of fixed effects, in the presence of unbalanced panel models; conversely, when balanced models are available, Durbin-Watson test is more appropriate (Baltagi, 2008). There are no Baltagi LBI critical values computed so it is not be possible to draw inferences from this test. However, stata reports the Durbin-Watson test statistics, which in both models suggest that the t-statistics are outside the range of the lower and upper bounds of the critical values, indicating a positive autocorrelation of the residuals. Further, the estimated ρ coefficient which is statistically different from zero in both models (ρ = 0.428 in precautionary; ρ = 0.446 in involuntary model, Appendix 4.2), also supports the rejection of the null hypothesis. Therefore, the common factor restrictions test, the F-test suggesting that fixed effects are significant, the Durbin-Watson and the estimated ρ coefficient, altogether suggest that the fixed effects model conditional on the AR(1) residuals is the correct model specification for out model. [90] In the precautionary model, the test statistics is lower than the critical value for the LBI test, suggesting that we do not reject the null of the serially independent residuals (Appendix 4.2). The LBI test diagnostic from the involuntary model is less straightforward. At 5% level of significance, the LBI t-value of 1.87 is in the middle zone of the lower and upper bonds of critical values of 1.84 and 1.89, suggesting the importance of taking into account the unobserved dynamics in the modelling strategy. Further, the robust standard errors were used to account for heteroskedasticity when conducting statistical inference. These standard errors are robust to the conditionality of the estimates of slope coefficients β on the autoregressive coefficient ρ, and vice versa.

In table 4.2 the results from the panel data fixed effects are reported for precautionary excess liquidity in panel a) and involuntary excess liquidity with linear relationship between excess liquidity and lending interest rate in panel b) and involuntary excess liquidity with nonlinear relationship between excess liquidity and lending interest rate in panel c). With regard to precautionary excess liquidity, only the deposit interest rate and growth expectations have turned out as statistically significant, at 10 respectively five percent levels of significance. An increase in the deposits interest rate by one percentage point, on average, would decrease the excess liquidity ratio by 0.4 percentage point, ceteris paribus. This finding is not in line with theory, since the interest on deposits represents a banks' cost towards their customers, and when these costs rise, banks are expected to accumulate greater excess liquidity. However, one explanation may be that in order for banks to pay higher interest on their customers' deposits, they may use their excess liquidity funds, this way reducing the level of excess

[90] Baltagi and Li (1991) applied the Prais-Winstten transformation matrix, to transform the disturbances into serially uncorrelated classical errors (Baltagi, 2008, p. 92).

liquidity. The banks' expectation about future growth prospects are in line with the prior expectations, indicating that a positive growth rate in the following year, would motivate banks to be less cautious and use up more liquidity buffers for investment, ceteris paribus. Further, the magnitude in the next year's GDP is relatively large, suggesting that banks are willing to invest larger amounts of their excess liquidity when growth prospects are positive. This may be explained by the fact that most of banks in European transition countries rely on a simple banking model, where main activity consists on accepting deposits on one side and lending the economy on the other side, hence making these two indicators more manageable. The sign on the result of non-performing loans indicator is in line with expectations, as an increase in non-performing loans, on average is expected to induce banks to hold more excess reserves and lend less for precautionary reasons, other things being equal. The time variation in the model was accounted with year dummies. All year dummies have turned out as statistically insignificant, thus it is not relevant to further comment on them.

Table 4.2. The determinants of excess liquidity

Precautionary EL		**Involuntary EL**		**Involuntary EL: iloans non-linear**	
Variables	Coefficients	Variables	Coefficients	Variables	Coefficients
rrr	0.209	*depgdp*	6.703**	*depgdp*	6.112*
	(0.493)		(3.318)		(3.258)
ideposits	-0.383*	*iloans*	0.227	*iloans*	-2.867*
	(0.219)		(0.516)		(1.600)
equity	-0.433	*credgdp*	-4.434	*iloans2*	0.130**
	(0.619)		(2.814)		(0.064)
rgdp1	-1.901**	*legal*	-9.665	*credgdp*	-3.578
	(0.835)		(10.960)		(2.785)
volcred	-2.247	*fdig*	0.686*	*legal*	-11.35
	(5.527)		(0.359)		(10.750)
voldep	-3.833			*fdig*	0.798**
	(2.581)				(0.356)
volgdp	1.902	*d2005*	9158	*d2005*	3616
	(2.524)		(22135)		(22864)
npl	0.087	*d2006*	4218	*d2006*	1661
	(0.190)		(10178)		(10444)
d2004	-2.735	*d2007*	1931	*d2007*	760.6
	(14.111)		(4649)		(4740)
d2005	19.138	*d2008*	867.6	*d2008*	336.9
	(15.047)		(2099)		(2127)
d2006	11.664	*d2009*	375	*d2009*	141.8
	(15.595)		(916.6)		(923.3)
d2007	13.257	*d2010*	151.4	*d2010*	57.21
	(15.772)		(369.8)		(370.5)
d2008	2.155	*d2011*	48.05	*d2011*	18.09
	(17.848)		(117.1)		(116.8)
d2009	14.605	*d2012*	-	*d2012*	-
	(15.736)				
d2010	20.720	*Constant*	113.3***	*Constant*	113.9***
	(15.787)		(37.73)		(36.988)
d2011	14.121				
	(16.096)				
d2012	-17.959				
	(15.737)				
Constant	0.557				
	(10.249)				
Observations	100		119		119
Number of banks	39		40		40

Note: ***Significant at 1% l.s., **Significant at 5% l.s., *Significant at 10% l.s.

Source: Author's own calculations performed in Stata 11.

The estimated coefficients on involuntary excess liquidity in panel b) tell a more consistent story in terms of coefficients sign. Deposits and loans represent the main relations with respect to excess liquidity, as deposits on one hand increase the excess reserves and on the other loans decrease them. Deposits are the main source of excess liquidity that commercial banks hold either in their accounts either in central banks. As expected, deposits to GDP are highly significant with a relatively large magnitude on the coefficient.[91] Our results suggest that, on average, an increase of one percentage point in deposits to GDP ratio would increase involuntary excess liquidity to assets by 6.1 percentage points, ceteris paribus. Referring back to section 4.2, this raises the question why do banks continue to accept deposits that they not necessarily need nor have capacity to invest.[92] Private sector credit to GDP, is significant at 5 percent level of significance. On average, an increase of one percentage point in credit to GDP ratio, would decrease involuntary excess liquidity by 7.4 percentage point. The FDI to GDP ratio turned out as statistically significant at the 10 percent level of significance and has the anticipated sign. The result suggests that an increase of one percentage point of FDI to GDP ratio, on average, would increase excess liquidity by 0.5 percentage point. The relatively small coefficient magnitude on FDI indicator may be explained by the fact that FDIs may take different forms and not all FDIs circulate via the banking system. The rest of the variables have turned out as statistically insignificant, but the coefficient signs are consistent with expectations. For example, the coefficient on loan interest rate suggests a positive relationship with involuntary excess liquidity. The strength of legal rights in collateral execution and bankruptcy laws was expected to be one of the most important factors with regard to market and institutional conditions. Nevertheless this variable has turned out as statistically insignificant. The reason may be that the measure of the strength of legal rights is an index variable with a relatively short time series (2007 - 2012) with very little within group variation. Therefore, as the fixed effects estimator mainly relies on the within group variation, it can render relatively time-invariant variables as insignificant, even though they are economically significant (Buddelmeyer et al., 2008).

Finally, the results in panel c) are quite robust with the estimated results in panel b) regarding coefficient signs, magnitude and statistical significance, thus the economic interpretation is the same. However, the non-linear relationship between excess liquidity and lending interest

[91] It should be noted that in a bank balance sheet deposits are proportionally much larger than excess liquidity funds, thus a small increase on deposits, may on average lead to a considerable increase in excess liquidity.
[92] Due to data limitations, we could not distinguish between private sector and government deposits to make inferences about which group of depositors cause higher excess liquidity.

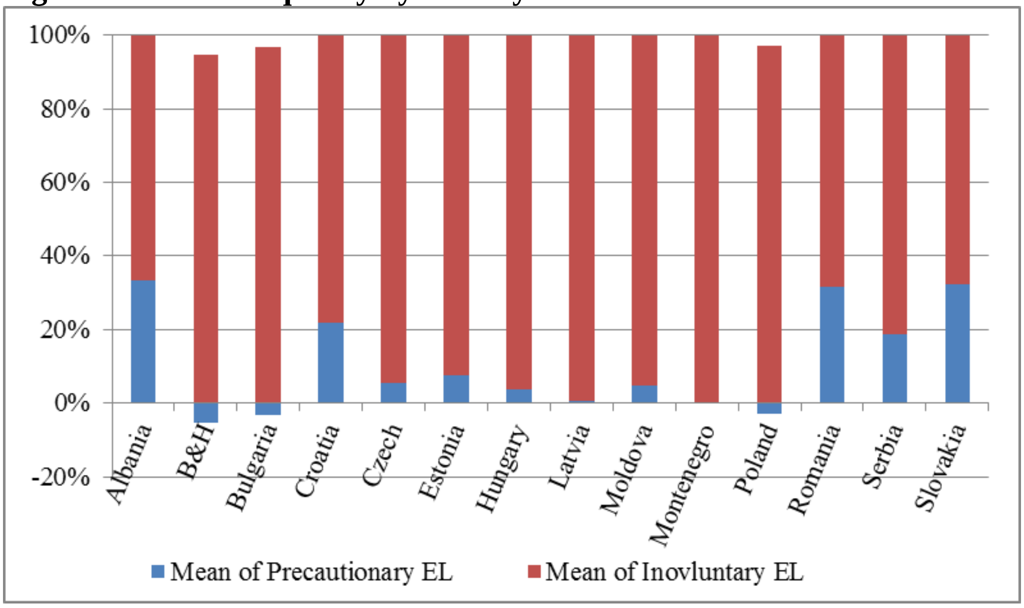
rate is significant in the panel c). This suggests that, even though an increase in loan interest rate is supposed to increase banks' expected returns and thus induce them to use up more excess reserves for lending, after some point, when interest rates are relatively high, the loan supply curve may become backwardly-bending, as the asymmetric information that exist between banks and borrowers creates barriers in the intermediation process. Therefore, contrary to the linear relationship between excess liquidity and lending interest rate proclaimed in orthodox theory, the findings for our sample model suggest that credit rationing theory in European transition economies' banks hold.

Having said that, it should be noted that both FE AR(1) models also controlled for country dummies, however stata dropped them all due to collinearity. This because the FE estimator estimates only the within group variation, and the time-invariant estimates are differenced out, but controlled for. Thus in the cases of small samples with low between group-variation, like is our cas, the FE estimator cannot generate efficient estimates, because the low variability makes it difficult to identify the effect. Even though FE estimator has many advantages, this represents a drawback.

4.4. The estimated portion of precautionary versus involuntary excess liquidity

After estimating and interpreting the results from the two FE AR(1) models, the next step is estimating the level of precautionary versus involuntary excess liquidity in banks in our sample. Even though the dependent variable – excess liquidity – is an observable and measurable portion of liquid assets, the portion of precautionary and involuntary is unobservable and needs to be generated from the model. The two original studies that generated the model to disaggregate the two portions (Agenor et al., 2004 and Saxegaard, 2006) used macro-level time series data so they could simulate dynamic predictions post-regression and plot the generated series of precautionary and involuntary levels with timeline plots. Given that in our models we utilise bank-level panel data with a high ratio of missing values, it was technically not possible to generate the dynamic simulations. Thus instead we have used the following procedure: a) the first step was estimating equations (4.16) and (4.17) as presented in table 4.2; b) next the fitted values of the models were generated (static predictions from the estimated models in post-regression), representing the precautionary and involuntary excess liquidity , respectively; and c) given the use of the panel data, instead of plotting the series in a timeline chart, so that the trend could be observed, the estimated series (precautionary and involuntary excess liquidity) were plotted in a bar chart. Before doing so, the estimated levels of precautionary and involuntary motives were aggregated from bank-level into a macro-level series. Further, having in mind the high ratio of missing values at the beginning of the series, it would also not be sensible to plot the generated series in a timeline plot, because even though the level of excess liquidity would be rising with time, in our case, the level of excess liquidity would be rising only because in the more recent years the number of the banks reporting has increased. However, changing the structure of the reporting of the generated series from other studies does not affect the implications of the results, the purpose of this model is not to measure the exact level of precautionary and involuntary motive for holding excess liquidity; it is rather to draw inferences about which motive prevails. The estimated series were generated as were aggregated in two ways, by countries (figure 4.3) and by years (figure 4.4). Afterwards, the aggregated figures were split as percentages of the total 100 percent of excess liquidity.

Figure 4.3. Excess liquidity by country

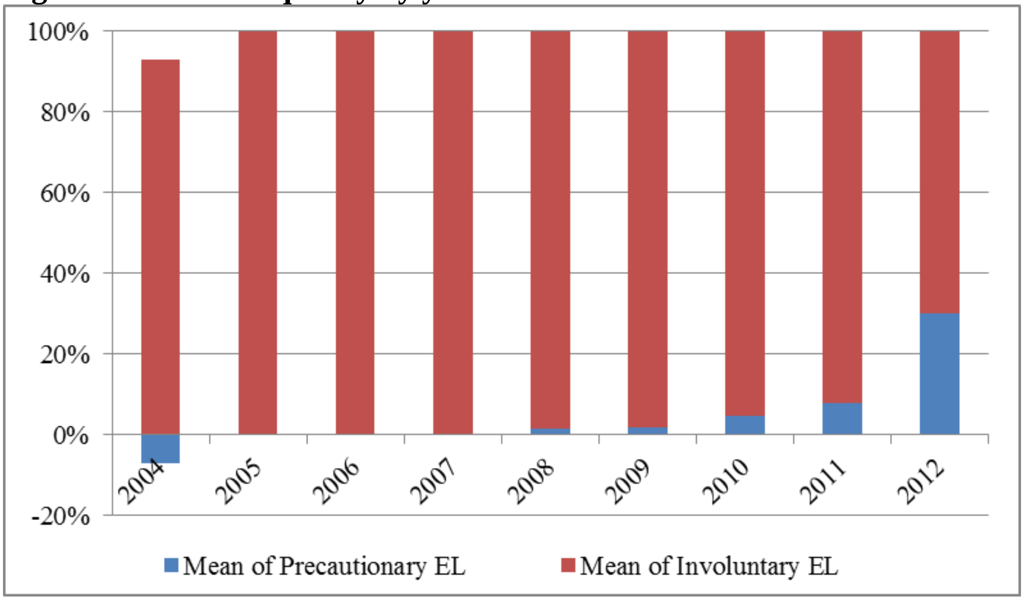


As it can be seen from the figure 4.3, in most of the countries in our sample, it is the involuntary motive for building the excess liquidity that prevails.[93]

[93] Another Baltic country, Lithuania was intended to be in the model, however due to the very high rate of the missing values in the dataset, it was removed from the model.

Bosnia and Herzegovina, Bulgaria and Poland represent a different case, where negative values of precautionary excess liquidity were generated.[94] This result indicates that on average of the period estimated, banks in these countries were continuously accumulating involuntary excess liquidity on one side. Albania, Romania and Slovakia are some of the countries with the highest portion of precautionary excess liquidity, yet even here does not exceed 30 percent of the total excess liquidity, thus the involuntary motive again prevails.

Figure 4.4. Excess liquidity by year



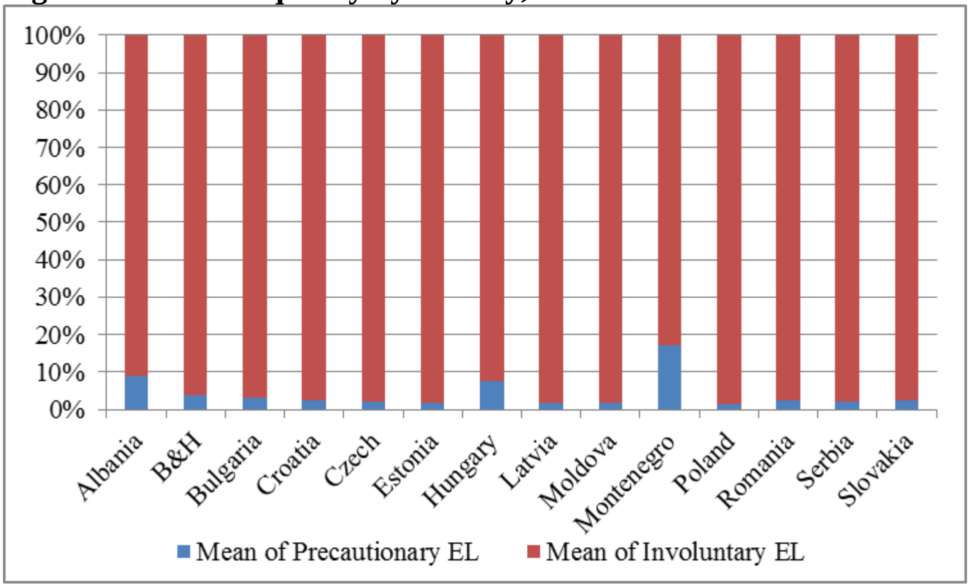
We now turn to the interpretation of the precautionary and involuntary excess liquidity over time. Even though the series in the dataset for all the banks start at 1998, years 1998 to 2003 are not presented in the plot. This is because the response rate of the banks reporting in Bankscope was very low, thus the generated results were not necessarily sensible. Figure 4.4 suggests that up to 2007, all the banks in the sample were clearly driven by involuntary motive to hold excess liquidity, meaning that if the capacity absorption for loans was higher, these banks would probably lend more and hold less excess liquidity. From 2008, which coincides with the commencement of the GFC, the results suggest that banks started to become more cautious and start to increase their precautionary excess liquidity. Nevertheless, even during these years, the involuntary motive prevails considerably, most probably due to depressed demand in a recession period in most of the European transition countries at that

[94] It should be noted that negative values of the excess liquidity (dependent variable) are present in the series, especially in the late 1990s and beginning of 2000s. The negative values represent the cases when banks used up a portion of mandatory liquid assets and therefore were considered to keep below the required levels of liquidity either in the central banks or above mandatory liquidity ratio.

time. The euro-crisis in 2010 and sovereign debt crisis that followed contributed to increasing the uncertainties in the banking systems, amongst others, by slowing down the issuing of new loans and new investments.

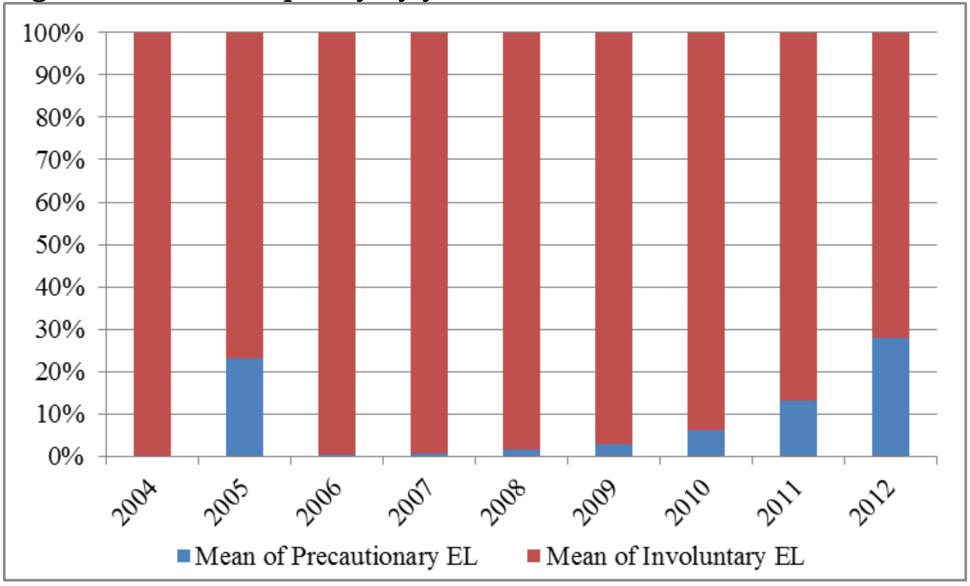
In the figures 4.5 and 4.6 the same disaggregation procedure was followed as in the two figures above, except that the estimated model also includes the non-linear relationship between excess liquidity and lending interest rate (see table 4.2 panel c)).

Figure 4.5 Excess liquidity by country, i^L non-linear



When comparing figures 4.3 with 4.5 and 4.4 with 4.6, the inclusion of the non-linear lending interest rate did not significantly change the disaggregated outcome. In both cases, the involuntary motive for holding excess reserves clearly dominates in all the cases, either by country or by year.

Figure 4.6. Excess liquidity by year, i^L non-linear



The concluding remarks of the chapter and possible policy implications will follow in the next section.

4.5. Conclusion

In this chapter we have introduced a new topic of the thesis, that of excess liquidity. Having found out that conceptualising excess liquidity is not straightforward matter in the literature, the first part of the chapter has given a considerable thought to its definition. In this line of arguments, two different sets of excess liquidity definitions were considered: one that only considers the above statutory reserve held at central bank and the other that considers liquidity in excess of mandatory requirements. Given that excess liquidity in our thesis is investigated as an underutilised resource and providing the potential for more lending in European transition economies, a definition that considers both parts of excess liquidity was considered to be more appropriate.

Alternative theoretical treatments of their determinants of excess liquidity were reviewed along with the limited previous attempts to quantify those determinants. It was argued that in designing an appropriate empirical model the key characteristics of European transition economies would need to be recognised. The model estimated was mainly based on the Saxegaard (2006) approach, which distinguishes between determinants of precautionary and involuntary motive for holding excess liquidity in two equations. The results suggested that the cost of liquidity funds as approximated by deposit interest rate and growth prospects are the most important factors inducing banks in European transition economies to hold excess liquidity for precautionary reasons. However, as earlier discussed in this chapter, this part of excess liquidity is part of a profit-maximising behaviour of the banks and cannot be treated as 'excess' from the economy's perspective. On the other side, the deposits held at banks, credit to the private sector as well as FDIs were the main determinants driving the accumulation of banks' involuntary excess liquidity.

In the second part of the estimations, excess liquidity was further disaggregated between the precautionary and involuntary motives. Both, by country and by year, the results clearly suggested that the involuntary motive prevails in inducing banks to accumulate excess reserves.

To conclude, the finding that most of the excess liquidity is involuntarily held in banks provides compelling reason to suggest that banks in European transition economies hold non-remunerated excess reserves that do not provide a convenience return. Some of the factors causing banks to involuntarily hold excess reserves are factors outside the banks' control, such as the lack of alternative savings vehicle other than banks, the lack of choice of instruments to invest other than deposits, asymmetric information, lack of absorption capacity from economy for new loans etc.

After investigating excess liquidity and its determinants, in the following chapter, the relationship between output gap and excess liquidity will be examined.

Chapter 5. The relationship between output gap and excess liquidity

5.1. Introduction
5.2. Theoretical relationship between output gap and excess liquidity
5.3. Endogeneity issue between output gap and excess liquidity
5.4 Data issues
5.5. Unit root tests for the SUR model
5.6. Methodology
  5.6.1. Single-equation approach: FGLS and PCSE
  5.6.2. System-equation approach: SUR model
  5.6.3. The SUR method in the context of output gap and excess liquidity framework
5.7. Interpretation of diagnostics
5.8. Interpretation of results
5.9. Conclusion

5.1. Introduction

As argued in chapter 1 and established in chapter 4 of this research project, the European transition countries have continuously been accumulating excess reserves despite the opportunities to invest in seemingly profitable projects (e.g. loans) in the economy. After addressing one of the project's key research questions 'why do banks hold excess liquidity in face of profitable loans' in the chapter 4, this chapter aims to address another key research question, that of 'does excess liquidity have an impact on the output gap'. Compared to the large set of literature investigating the output gap and the few investigating the determinants of excess liquidity in particular, this chapter is novel in two respects. Firstly, a previously neglected conceptual relationship between the output gap and excess liquidity is investigated and secondly, a new econometric method is applied to account for potential endogeneity between the two. In the theoretical part of this chapter, section 5.2, both concepts, that of excess liquidity and the output gap, will be reintroduced and a further analysis of their possible relationship will follow. Given that excess liquidity, if translated into further loans for the economy is believed to impact on growth and production capacity, we will analyse its impact in the context of the output gap. Further, in section 5.3, the possible causal relationship and endogeneity issue between the two will be examined. The theoretical discussion will be followed by examination of the data issues in section 5.4, followed by unit root tests of all variables in section 5.5. The methodology used for empirical analysis is discussed section 5.6. The estimated results and their interpretation is presented in section 5.7. Section 5.8 presents additional estimated model as a robustness check to the preferred specified model. Section 5.9 presents the concluding remarks.

5.2. The theory relating the output gap and excess liquidity

Building a theoretical relationship between the two important macroeconomic variables - the output gap and excess liquidity- proves to be quite challenging for various reasons. Firstly, the studies that have investigated potential output and the output gap are mainly concerned with the estimation techniques used to generate these two unobserved variables and do not relate to excess liquidity. Secondly, the emphasis of these studies on potential output and hence the corresponding output gap is on their relevance for the conduct of monetary and fiscal policy (Krugman, 1998; Eggertson and Woodford 2003; Agenor et al., 2004; Polleit and Gelldesmeier, 2005 and Saxegaard, 2006). Thirdly, as also previously noted, excess liquidity is treated quite vaguely in the literature, both in terms of its definition and as a general phenomenon in the transition countries and when it is considered it is mainly regarded in the light of monetary policy effectiveness. Fourthly, the importance of the relationship between business cycles and financial cycles in macroeconomic systems became a major issue of concern only after the onset of the Global Financial Crisis (GFC). Finally, and most importantly, because to the best of our knowledge this specific relationship – output gap versus excess liquidity – has not been previously investigated, especially in the transition context.[95] It was only mentioned briefly in Saxegaard's (2006) analysis, where our relationship of interest is presented the other way around (that is excess liquidity being the dependent and the output gap one of the control variables) and it did not represent the main relationship of that study and used but later dropped by Agenor et al. (2004).[96]

As discussed in chapter 2, potential output is the level of output produced with fully employed factors of production without generating inflationary pressure, whereas the output gap represents the difference between the actual and the potential output. The size of the output gap indicates the position of the economy in the business cycle. Moreover, as the conventional (Neo-Keynesian) theory proclaims, the output gap emerges as a key indicator of inflationary pressure; for instance, a positive output gap may be seen as an inflationary signal. This is because the mainstream theory has always modelled the output gap in association with conventional relationships such as the Phillips' curve or Okun's Law. Asserting that potential output is always associated with stable inflation is too simplistic and not necessary the only case. As Biggs and Mayer (2013) argue, the relatively high growth rate accompanied by a relatively low volatility, especially in developed countries, was not necessarily a result of technological advancement and central banks "who had conquered the monetary policy through the concept of inflation targeting". The booming period during the first part of 2000s was rather the impact of the upswing of the credit cycle, which culminated in 2007. The recent global financial crisis has shown that imbalances (such as financial imbalances and housing market imbalances) can build up even in a relatively stable inflationary environment and ultimately lead to severe disruptions in the economy (Furlanetto et al., 2014). Therefore, sustainable output and non-inflationary output need not coincide. Biggs and Mayer (2013) argue that the conventional approach of analysing the relationship between the two was through comparing credit growth with GDP growth or spending growth. This is a stock/flow

[95] Given the lack of the theory specifically addressing the relationship between output gap and excess liquidity, the arguments relating the two were derived from the growth and finance nexus literature.
[96] Saxegaard (2006) uses the five-year moving average of the standard deviation of the output gap as one of the risk variables determining the precautionary motive (equation) for holding excess reserves, as well as output gap (as a percentage deviation of output away from a quadratic trend), in the same equation.

error – growth in the stock of credit is compared with growth in the flow of GDP. As a result, the relationship between the two is perceived to be weak. Instead, the authors argue that the business cycle should be compared with the flow of credit (or new loans) or the growth rate with changes in the flow of credit. [97]

Currently the Euroarea may serves as an example of a low-inflationary environment with low loan interest rates and the ECB's discount window reaching its historical low rates (0.25% to 0.15% and even negative rates recently), yet many of the member countries are still struggling with recession caused by the financial crisis. As a consequence of lower intermediary activity accompanied by depressed demand, banks in these countries have been operating with excess liquidity since late 2008 (ECB, January Monthly Bulletin, 2014).

In addition to the supply of new credit as an important financial variable that can well explain the output gap in an economy (Biggs and Mayer, 2013) our focus is on an unexplained and (partially) unutilised part of credit, that of excess liquidity. The excess liquidity is treated as a potential (unutilised) credit for the private sector of the real economy. Investigating the role of excess liquidity in this context is considered as important, because the endowment of potential resources in the economy is expected to create output. Therefore, the important research question to be addressed in this chapter is 'does excess liquidity impact the output gap'?

As discussed in chapter 2, the output gap can be positive or negative, depending on the position of the business cycle. A positive output gap occurs when actual output is more than full capacity output. This happens when demand is very high and to meet that demand, factories and workers operate far above their most efficient capacity. In the transition context, however, where the production capacity is frequently considered to be low even when demand pressures are high, the overheating phase of the cycle may also be reflected via higher imports and hence higher trade deficits. On the other hand, a negative output gap means that there is spare capacity or a slack in the economy due to weak demand. In a transition context, via higher trade deficits, a negative output gap may be present. Under these circumstances and in the presence of excess liquidity banks can shift the composition of

[97] Borio (2012) states that there is no consensus yet on the definition of a financial cycle: 'The term will denote self-reinforcing interactions between perceptions of value and risk, attitudes towards risk and financing constraints, which translate into booms followed by busts. These interactions can amplify economic fluctuations and possibly lead to serious financial distress and economic dislocations. This analytical definition is closely tied to the increasingly popular concept of the "pro-cyclicality" of the financial system' (Borio, 2012, p.2).

savings toward capital, causing intermediation to be more focused on growth promoting expenditures (Bencivenga and Smith, 1991). Thus, the banking finance can have a more significant impact in growth via resource allocation than through pure capital accumulation (Beck et al., 2012).

On the other side, the presence of excess liquidity in the banking sector, given that it reduces banking efficiency and profitability and holds back further lending, can be considered as a deadweight loss for these economies (see chapter 4). Thus, excess liquidity in our study is treated as (partially) unutilised funds available for further lending in the economy.[98] Given that excess liquidity is treated as 'excess' from the economy's perspective and not necessarily from the banks' perspective, this part of the banking funds could be viewed as lowering the level of credit available in European transition economies. As discussed in chapter 1, a desire to see banks lend more to the private sector of the economy is usually a familiar topic in transition context. This is because banks in European transition countries are considered to be the main the financial resource of the households and enterprises, given the lack of capital markets. As argued in the chapter 1, banks in these countries seem to supply less credit to their economies compared to the banks in developed countries. Low ratios of credit to GDP may also reflect the lack of funding capacity of banks, however, given that most of the banks in these countries fund their loans from domestic deposits (see chapter 1), this does not seem the case. Most of these banks, with the exception of foreign banks which initially fund their loans from parent banks, do not take the risk of raising funds externally, but mainly fund the loans with domestic deposits.[99] The presence of excess liquidity in the commercial banks of transition economies serves as evidence that banks in these countries may have the capacity to further extend their lending. Given the low credit ratios together with the presence of excess liquidity, we can argue that the demand for loans in transition economies is not being met (compared to the more developed countries). Thus, the accumulation of excess liquidity in transition economies is expected to have two effects via the same causal mechanism:

[98] A detailed discussion on the definition and determinants of excess liquidity in the commercial banks of transition economies is presented in chapter 4.
[99] Besides using excess reserves, one may argue that banks can also utilise external funding or part bank's credit lines if they wish to extend the lending. However, in the transition economies where the financial market is mainly bank-based, domestic deposits comprise the main source of banks' funding. Compared to the banks in developed countries where they have access to cheaper funds in capital markets, the deposits of the banks in transition economies as a funding source are relatively more expensive. Thus, instead of borrowing elsewhere, we argue that banks would rather use up the readily available funds from their deposits, i.e. excess reserves.

- The level of lending is lower, thus as a consequence, growth of output capacity is likely to be constrained, as there is a capital shortage.

Thus, the argument for concentrating upon the output gap versus excess liquidity relationship lies in the proposition that in some transition economies output is capital constrained and excess liquidity can be high even when there is no output gap.

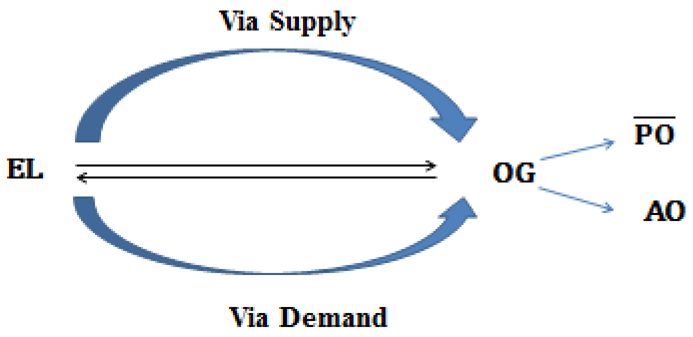
Although, normally when output is close to potential the increasing firms' demand for loans would be met by an increase in supply, in some economies credit is rationed and increased lending opportunities are not taken up. With efficient markets, when firms are approaching full capacity utilisation, in macro-level the actual output is approaching potential output; in this situation firms would increase the demand for loans for investment finance and banks would respond by increasing the provision of investment finance. However, when markets are imperfect, credit rationing happens and the credit opportunities are not taken up by banks (Stieglitz and Weiss, 1981). Imperfect markets with considerable factors impacting the information asymmetry appear to be a structural problem of transition economies (World Bank, 2013). Banks hoard liquidity partly as they may not view these lending opportunities in boom phases as likely to be profitable. If banks slow down credit growth and increase the levels of excess liquidity, than the actual output would grow slowly, and in longer run, potential output would be lower. Furthermore, if demand is growing more slowly, than the actual output would grow at a slower pace, too. In effect, in those countries where less loans are issued, presumably excess reserves will build-up more in the banks, which in turn would also imply lower demand potential growth. In the aftermath of the financial crisis, transforming excess liquidity into further credit may be even more important, since they can be also used as an instrument to smooth the recession phase, i.e. behave counter-cyclically. In this case, Biggs and Mayer (2013) argue that all that it is required after a ‘credit crisis’ for a recovery in demand growth is that the *new borrowing* (flow of credit) rises. If properly compared, that is comparing the flow of credit with the output gap, instead of the stock of credit (as usually mistakenly taken in studies) then the authors argue that the relationship between credit and business cycles is strong. It is indeed the *change in new borrowing* (credit impulse) that is perfectly and positively correlated with the demand growth. The authors find that the new loans cycle from the banks comprise very close information regarding the business cycle, so that new loans can be used as a proxy of the business cycle.[100] This way, the spending in a particular period will increase if new borrowing takes place in that period, leading into an increase in demand.

Finally, the excess liquidity thus transformed is expected to enhance potential demand and increase production capacity. If this is the case, then savings held in the banking system in the form of excess liquidity matter for the economy, and further lending would help in maximising the growth benefits of finance. It should be noted that rather than the size of the output gap, what's important is whether the potential output is increasing. A negative output gap is likely to constrain the growth of potential output. Eliminating a negative output gap is beneficial to long-run growth, since that creates incentives to firms to expand capacity.

On these arguments alone and especially the Biggs and Mayer's one, there may be no relationship between excess liquidity and output gap, because, excess liquidity, in addition to its expected impact on the output gap, may also impact the other two components of output, actual output (AO) and potential output (PO), via demand and supply channels (Figure 5.1.). This is because, an increase of potential credit via a reduction of excess liquidity may increase the economic activity by increasing income and consumption; however, the part of the excess liquidity transformed into capital investment loans is expected to increase the potential output also, via the supply side channel. Furthermore, if the impact of excess liquidity on actual and potential output is symmetric, then both are expected to move in the same direction: less excess liquidity and more potential loans would increase both, whereas more excess liquidity and consequently less potential credit for the economy is expected to reduce both at the same amount. Nevertheless, compared to the actual output, the potential output is usually relatively stable or fluctuates relatively slow. Thus, we abstract the supply side impact on potential output and assume the potential output as fixed. However, empirically we expect a relationship between excess liquidity and output gap as well as actual output, given that these two fluctuate relatively more in short-run.

[100] Biggs and Mayer (2013) state that the credit growth was negative in 2010 in OECD countries, but because credit growth was less negative than in 2009, the credit impulse was positive and helped in boosting demand.

Figure 5.1. The impact of excess liquidity on output gap



However, just because an increase in new loans would have a causal effect on demand growth and consequently on the output gap, the same story can also be told the other way around, i.e. the causality may run from output gap to the excess liquidity. This is because, just because two factors are correlated, does not mean that one causes the other. A correlation simply implies that two factors are bound together, but is not much indicative about the direction of that relationship. If this is the case, than it is possible that the excess liquidity is causing output gap, just as it is possible that output gap is causing excess liquidity, or it may simply be the case that both are being shifted from other third party factors. Furthermore, following Biggs and Mayer (2013) argument and our empirical findings later on (see section 5.6.1) it appears that the output gap and excess liquidity are rather in a correlation relationship.

Therefore, although at the beginning we stated that excess liquidity is expected to impact the output gap (causality running from excess liquidity to the output gap), the relationship between the two may in fact be correlation, because they both may be measuring the same phenomenon, that is the general state of the economy. Thus, considering the theoretical background and other possible relationships between the two, we reformulate the initial research question form 'does excess liquidity impact output gap' to 'what is the relationship between output gap and excess liquidity'.

The possible relationships between output gap and excess liquidity will be discussed in the following section.

5.3. Endogeneity between output gap and excess liquidity

Uncovering the causal relations between economic variables and distinguishing them from associational relationships is one the most important tasks in economics (Zellner, 1979). This is because only causal relations are useful for policy advice, because they contain the reaction of the economic variables of interest to policy interventions (Lechner, 2011). Therefore, understanding the nexus between financial and business fluctuations is one of the important relationships for further guidance of the policymakers. The relationship of interest in our case is the one between output gap and excess liquidity. Nevertheless, given the lack of the specific theoretical and empirical guidance on our specific relationship, the discussion will start from a broader perspective, that of growth and finance nexus.

In general terms, the literature on the relationship between growth and financial development is vast and inconclusive. The question of whether financial development causes economic growth or economic activity causes growth of the financial sector is one of the most debated issues of neo-liberal development research. An extensive part of the literature investigating this relationship, especially prominent authors from endogenous growth literature, has supported the view that financial development proceeds growth and that financial development is strongly associated with current and future rates of economic growth (Greenwood and Jovanovic, 1990; Bencivenga and Smith, 1991; King and Levine, 1993, Greenwood and Smith, 1997). Even though these studies do not explicitly state the way the causality runs, generally they argue in favour of financial development being a promoter of economic growth (Bencivenga and Smith, 1991; King and Levine, 1993; Ross, Loyaza and Beck, 1999; Beck, 2012). Furthermore, it has been argued that the financial development is not only important in fostering real economic growth, but also in dampening the volatility of the growth process. However, financial development can increase growth but can also increase the volatility by increasing new sources of volatility. Yet, as observed from the recent financial crisis, it may also be argued that the financial sector not only can promote growth in good times, but as happened since in the onset of the financial crisis, it can also cause severe recessions.

On the other front, some economists raised their scepticism about the role of finance in economic development. Robinson (1963) for example, has suggested that financial development follows economic growth. Likewise, Lucas (1988) concluded, "the importance of financial markets is badly overstressed". In other words, even though these two may be strongly and positively correlated with each other, it is actually the pace of economic growth that paves the road for financial development that follows (Boulila and Trabelsi, 2004; Chakraborty, 2013).[101] In a more recent study, which investigates the impact of credit on growth after the financial crisis in 39 economies where the crisis was preceded by credit booms, Takats and Upper (2013) find that, from the lowest point of recession, changes in bank credit did not correlate with growth during the first two years of recovery. This study may be relevant in our research case, given that many European transition economies had experienced credit booms prior to the crisis. A more severe case may be Estonia and Czech Republic, whereas in Kosovo (even though did not experience recession post financial crisis) the credit was growing with double figures.

In addition to studies confirming a causal relationship between growth and finance, a third part of the literature claims that this relationship is likely to be more complex than suggested in many earlier studies. For example, Demetriades and Hussein (1996) find evidence of bi-directionality and some evidence of reverse causation. They also highlight the dangers of statistical inference based on cross-section country studies, which implicitly treat different economies as homogeneous entities, as the results tend to differ depending on the level of financial development. Peia and Roszbach (2011) argue that the causality patterns depend on whether countries' financial development stems from the stock market or the banking sector. In more developed countries with developed stock markets, the financial sector tends to cause growth, while in the bank-based countries (which is usually the financial structure of the transition economies) a reverse or bi-directional causality is present between banking sector development and output growth.

[101] Even though the literature on the causality between growth and finance is largely inconclusive, it appears that there is a publication bias towards the papers that favour that causality runs from financial development to growth.

However, it should be noted that the previously discussed relationship between financial development and economic growth has largely been discussed in the long-run context (as this is a long-run relationship theory). In the context of our research, it would be of interest investigating whether the two, financial development and economic growth, have a causal relationship or move together. However, given the data constraint, we will not be able to empirically investigate this sort of relationship for the countries in our sample over the long-run, thus the previously discussed theory on growth and finance does not directly relate to our specific research. There have been very few studies looking at this relationship over the business cycle for short-run relationship. Our research interest in relating the business cycle with the financial factors lies somewhere in between the short and long-run. Furthermore, as stated earlier, the theoretical and empirical guidance to investigate this relationship is almost inexistent, thus one should rely on the few studies investigating the relationship between output gap and credit. These studies model the relationship between output gap and credit are few and mainly rely on descriptive statistics to illustrate their strong positive correlation (Borio, 2012; Biggs and Mayer, 2013) so alternative models are still emerging. For example, Furlanetto et al., (2014) find that financial frictions (as measured by credit spreads using VAR method) in U.S. explain 21.56% of the volatilities in output. Moinescu (2012) investigates the impact of credit flow on the output gap in ten European Union member states (including the Czech Republic and Estonia) and finds that 15% of credit flow is reflected in the deviation of economic growth from its potential (using panel fixed effects method). These studies mainly argue that the financial factors impact the business cycle. Whereas, Biggs and Mayer (2013) build arguments based on descriptive statistics that the relationship between business cycle and credit is so strong that the flow of credit itself can easily approximate the business cycle itself. Nevertheless, none of these aftermath crisis studies consider the possibility of a reverse causality or a possible simultaneous relationship or any other source of endogeneity between the output gap and financial factors.

However, just as the business cycle may depend on money market behaviour, the money market behaviour may also reflect the a priori position of the business cycle. Thus, the causality between output gap and excess liquidity may run both ways. One side of the argument is that banks may credit ration borrowers, despite high demand the loans, thus releasing smaller portion of their funds at higher rates while accumulating more excess reserves. This way, banks may not be providing sufficient loans to stimulate further economic growth. If this is the case, the argument that the lack of credit is dampening growth and widening the output gap may stand, then the causality runs from finance to growth. On the other side, banks may argue that excess liquidity is simply an outcome of depressed economies, the lack of feasible projects to invest and low loan demand. If this is the case, then probably causality between the two runs the other way around.

Despite the fact that the causality may run one or the other way, a third possible relationship between output gap and excess liquidity is may the one where both cause each-other simultaneously. In other words, output gap (demand side)[102] and excess liquidity (supply side) may be jointly determined by other, underlying factors. For example, lower growth expectations lead to lower lending decisions and thus higher excess liquidity, meanwhile lower lending opportunities may cause lower levels of growth. This may arise due to growth expectations that banks have which may approximate for the loan repayment capabilities of the public, and influence banks' decisions if they should lend more or lend less). For example, when banks set up their annual plans and targets for the upcoming year, they usually base them on the expected economic prospects.[103] During the expansion phase of the business cycle, financial industry usually realizes profits and extends lending via cheaper loans and lower screening criteria. Meanwhile, the risk perceptions associated with the loan repayment is fairly low. The borrowers on the other side react to the cheaper loans by increasing their demand for credit. Given that the economy is experiencing a period of growth, the policymakers may also encourage the expansionary lending by banks. This way, by accumulating from one period to another, both business cycle and credit cycle (and less excess liquidity in this case) become pro-cyclical (Kastrati, 2012). Arguably, the opposite may apply in times of recession; banks become more reluctant to lend, so they shrink or slowdown the pace of credit and hence accumulate more excess reserves, this way deepening the recession further, as noticed after the financial crisis.[104] Overall, since we cannot expect

---

[102] Demand variables are usually proxied by the output gap, the unemployment gap or the rate of capacity utilization (Gordon, 1997).

[103] It should be noted that in this respect, contrary to the mainstream theory which investigates the GDP growth rate relation with the credit in an economy, in our case, it is rather business cycle prospects that should matter, i.e. the level output gap instead of the say next year's GDP. This is because banks in principle should be preoccupied with the absorption capacity of an economy, feasible projects to invest and sustainable demand for their loans. Having said that, even if the GDP growth of the next year is expected to be positive but the economy is near potential or above potential that would imply a fed up demand for loans. On the other side, if the GDP growth is fairly low or even expected to be negative, a below potential output (negative output gap) would imply capacity for new loans. In this respect, it should be the position of the business cycle i.e. the output gap that should matter more than the GDP (or GDP growth) itself.

[104] Calvo et al., (2006) argue that in times of recession or crisis, when a sudden stop or considerable slowdown of the credit emerges, the funds available, instead of being used as loans would be either redirected into other instruments like government bills and the rest would be accumulated simply as excess liquidity in (cont…)

the change in one to not also impact the change in the other, the overall relationship between output gap and excess liquidity should be seen as an entire system, in order to observe the feedback loops involved.

However, it should be noted that the adjustments of excess liquidity and the output gap do not happen instantaneously and adjustments may take time. For example, if policymakers induce banks to further extend lending (e.g. via lower discount rates or some other quantitative easing process), it might take considerable time for this credit to reach households and firms, due to asymmetric information and administrative procedures (Bernanke et al., 1999). Therefore, the transformation of excess liquidity into more loans for economy and finally its contribution to the output gap may result with some lag effect. Thus, besides the possibility that these two factors may co-determine each-other on the current system framework, past decisions on excess liquidity may impact the current level of output gap also, as well as future level of output gap may influence the current decisions on using up or accumulating excess liquidity. Therefore changes in the feedback loop of the two macroeconomic factors may reflect with some delays. This phenomenon may be attributed to the strong 'lag-effect' typical of the business cycle (Kalimani, 2009).

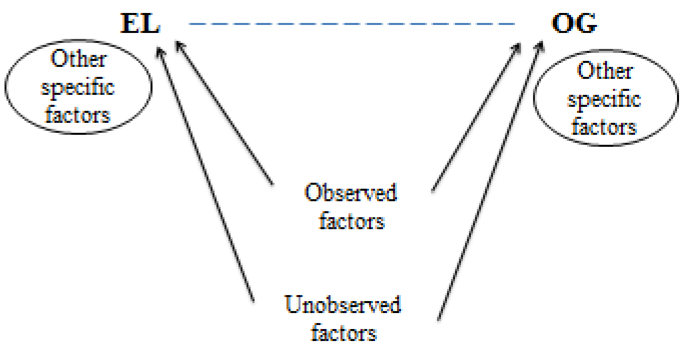
If reducing excess liquidity can contribute to raising the potential output and reduce the negative output gap in European transition economies, it is of great importance to investigate this relationship because it provides relevant information to the policy makers and regulatory institutions (e.g. central banks, macro-prudential institutions, ministries of finances etc.) to provide a favourable environment necessary for the growth of the financial sector (supply leading hypothesis). Furthermore, the study of this relationship may be ultimately important, given that financial development (especially in the transition context) typically increases the effectiveness of monetary policy (Biggs and Mayer, 2013), widens fiscal policy space and allows a greater choice of exchange rate regimes (IMF, 2012).

However, at this point, we can say that framing on initial research question may be misleading, because it directs us to investigate a causal relationship from excess liquidity to the output gap. At first, this implication seems reasonable, as based on the arguments

(…contd.) commercial banks resorts. Hence, once again excess liquidity may reduce the banks' profitability and efficiency, but more importantly (from the perspective of our research study) it would also imply a deadweight loss from the real economic perspective.

presented above excess liquidity accumulation appears to reduce economic activity. Yet, as we have seen further on, the story can also be told the other way around (Peia and Roszbach, 2011; Rachdi and Mbarek, 2011) that large output gap reduces the possibilities for lending. However, this does not necessarily imply that the alternative to excess liquidity causing output gap or output gap causing excess liquidity is simultaneity, i.e. both cause each-other. In the literature, the output gap is usually used as an indicator of the business cycle. Yet, Biggs and Mayer (20130 argue that the new flow of credit (or new loans) can easily approximate the business cycle itself. This suggests that excess liquidity seen as potential credit and output gap are different ways of looking at the same phenomenon, i.e. they both reflect the general state of the economic activity. In other words, because they may be co-determined, they are correlated. This view is consistent with the Biggs and Mayer (2013) finding that between business cycle and new borrowing "the correlation in excellent". If the relationship between output gap and excess liquidity is *correlation*, than this relationship suggests a different perspective on their causation and relationship. In this case, they both are outcomes of a wider system. In addition to being co-determined, output gap and excess liquidity may also be separately determined by other observed and unobserved influences (exogenous factors), which do not enter in the feedback cycle, as depicted in figure 5.2. In a transition context, some of these unobserved factors impacting both may be: informal economy, lack of law enforcement, corruption, banking reforms, etc. (Ease of Doing Business, 2013).

Figure 5.2. The relationship between output gap and excess liquidity



Being co-determined due to correlation is a radically different view of causation. This suggests that a change in excess liquidity does not happen because a change in output gap occurs, nor does it suggest that a change in output gap happens due to a change in excess liquidity. Revealing the sort of the relationship for these variables is of utmost importance for policy implications. From a policy perspective, changing excess liquidity will not change the output gap, and vice versa. This is because, other factors in the system would have to change to shift both (figure 5.2). If the other factors in the system (observed and unobserved) change, then both excess liquidity and the output gap will shift, but not because they are causing each other. Thus, in this type of relationship, even if policymakers induce banks to reduce excess liquidity to extend lending, it would lead to misallocation of resources.

To summarize, the inferences for further policy conduct essentially depend on the possible relationship between output gap and excess liquidity. If the relationship between the two is causal, then the policymakers can pursue banks to reduce excess liquidity and further extend the lending in favour of higher capacity growth and lower output gap. If the two factors are simultaneously determined by each-other, than changing one would automatically imply the change of the other, too. If however, output gap and excess liquidity are merely correlated and jointly determined by other (observed and unobserved) factors, then change in output gap and excess liquidity would not be possible without firstly shifting other factors (e.g. legal infrastructure) that impact both. In effect, this relationship would have to be reflected in an appropriate empirical framework.

Overall, as theory suggests, the relationship between output gap and excess liquidity is more complex from the uni- or even bi-directional causal one that was initially thought to be. Conceptually, the establishment of the relationship between our two variables of interest in a correlation perspective would also imply an empirical analysis appropriate to reflect this sort of the relationship. Therefore, in the following a discussion regarding the most appropriate approach to model this relationship will follow.

5.4. Data issues

The data period ranges from 2002Q1 to 2013Q4, with 48 observations per country, when the series has no missing values. The dataset is panel with three countries, namely the Czech Republic, Estonia and Kosovo.

A few series, especially for Kosovo (*$ideposits_t$, $iloans_t$, NPL_t, FDI_t,*) start later than the beginning of our dataset, 2002Q1. This is because official data compilation for specific series and sectors in Kosovo generally started at a later date, compared to other transition economies. Thus, while the data series for Czech and Estonia will start from 2002Q1, the dataset for Kosovo will start at 2004Q1. Given that the SUR modelling requires a balanced data (see sections 5.6.2 and 5.6.3), and in order to be able to utilise a larger dataset, the missing values of the series of some of the series at the beginning of the period are partially interpolated, using the cubic spline interpolation approach. The cubic spline approach is desirable in our case, since it allows for non-linearities in the series. Further, cubic spline interpolation avoids large changes in the series (compared to e.g. multiplicative approach which may generate very large values) from year to year and overall provides a smoother line. However, it should be pointed out that the interpolated values may not necessarily represent the actual values, so as it will be discussed and addressed later, a measurement error may be present in the data. The interpolated periods for each variable in Kosovo case are: interest rate on deposits and loans for 2002Q1 only, non-performing loans from 2002Q1 to 2005Q4, the FDI variable is interpolated from 2004Q1 to 2006Q4 (Table 5.1). The deposit and loan interest rate interpolation in Kosovo case may be acceptable given that the two series have been fluctuating around a similar average for ten years, thus are fairly predictable. As can be noticed in Table 5.4, the deposit interest rate starts with the actual value of 2.78 percent for the period 2004Q2, whereas the interpolated value one quarter before is 2.48 percent. On the other side, the loan interest rate also in Kosovo has also fluctuated around an average of 14.3 throughout the whole period (10 years), making the series quite predictable. The actual data for the loan interest rates in Kosovo in 2004Q2 start with 14.9 percent, whereas the interpolated value for the 2004Q1 is quite close, i.e. 14.85 percent. Thus, the interpolation of the Kosovo deposit and interest rate may be justified on the grounds that both have been relatively stable and predictable for a 10 year period. The NPL series has a longer range of missing values, therefore the data for 2004Q1 – 2005Q4) were imputed via the cubic spline interpolation approach.[105] However, since the beginning of the series with actual data (2.1 percent in 2006Q1) the NPL ratio in Kosovo has been increasing at a linear pace of around 1pp per year (in 2008 decreasing by 1pp), reaching 8.5 percent in 2013Q4. At the beginning of the series with the actual data in 2006Q1 the ratio of NPL is 2.1 percent. Before imputing

[105] The interpolation of the data represents a mechanical exercise, which fills in the gaps based on the actual values of the series. Imputation on the other side is applied when there are no data or a part of the series (e.g. beginning of the series as in our case) is fully missing. The imputation process goes beyond the sample period and estimates the missing periods by exploiting the correlation in the data.

the missing data for 2004Q1 – 2005Q4, the initial ratio at the beginning of the series was set to 1 percent. The relatively low rate of the NPLs in this period may be justified on two grounds: firstly because the transition period for Kosovo has started approximately in 2000 and secondly because at that time there were only two banks operating in the market, so the level of loans issued by then was quite low (3.3 million euro in 2000 or 0.1 percent of GDP, 25.9 million euro in 2001 or 0.9 percent of GDP and 86.5 million euro in 2003 or 3 percent of GDP and 232 million euro or 7.8 percent of GDP in 2003) and the time for the bad debt to accumulate was relatively short.

A detailed description of each data series is presented in table 5.1 below.

Table 5.1. Data description

Variable	Variable symbol	Description	Interpolated	Seasonally adjusted	Source
Excess Liquidity	*EL*	The ratio of excess liquidity (comprising above statutory reserves at the central bank+above mandatory liquidity ratio) over total assets. In percent.	No	Yes	Respective central banks.
Required Reserve Ratio	*RRR*	The reserve required ratio from the central bank to total deposits. In percent.	No	No	Respective central banks.
Deposit Interest Rate	i^D	Average interest rate on deposits. In percent.	For Kosovo (2004Q1)	No	Respective central banks.
Real Growth	*rgdp*	The real GDP growth rate compared to the previous quarter. In percent.	No	Yes	EUROSTAT, CBK
Volatility of Private Sector Credit	*VolC*	The five-quarter moving average of the standard deviation of the private sector credit, then divided by the five year moving average of the variable.	No	No	Own calculations*
Volatility of the Deposits	*VolD*	The five-quarter moving average of the standard deviation of the total deposits, then divided by the five year moving average of the variable.	No	No	Own calculations*
Volatility of GDP	*VolGDP*	The five-quarter moving average of the standard deviation of the real growth rate, then divided by the five year moving average of the variable.	No	No	Own calculations*
Non-Performing Loans	*NPL*	The ratio of of non-performing loans (delayed in repayment more than 180 days) over total loans. In percent.	For Kosovo 2004Q1-2005Q4	No	Respective central banks.
Equity	*Eq*	Total banking equity over total assets. In percent.	No	No	Respective central banks.
Deposits	*Dep*	Total deposits minus government deposits over GDP. In percent.	No	Yes	Respective central banks.
Government Deposits	*Govdep*	Government deposits (central + local) over GDP. In percent.	No	Yes	Respective central banks.
Credit	*Cred*	Total private sector credit over GDP. In percent.	No	Yes	Respective central banks.
Loan interest rate	i^L	Average interest rate on loans. In percent.	For Kosovo (2004Q1)	No	Respective central banks.
Rule of Law Estimate	*RLE*	Capturing perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.	Yes (2002 and 2013)	No	World Bank
Foreign Direct Investment	*FDI*	Foreign direct investment inflows over GDP. In percent.	For Kosovo: first 3 quarters of 2004-2008	Yes	Respective central banks.
Output Gap	*OG*	Output Gap over GDP. In percent.	No	Yes	Own calculations
Credit flow	*CredF*	Credit flow over GDP. In percent.	No	Yes	Respective central banks.
Eurozone Output Gap	*OGEU*	a) OECD: Output Gap over Potential Output b) HP filter: Output Gap over GDP. In percent.	a) Yes b) No	Yes	a) OECD b) Own calculations
Government Expenditure	*Govexp*	Total government expenditure over GDP. In percent.	No	Yes	EUROSTAT (Czech and Estonia), Ministry of Finance (Kosovo)

*Note: The calculation of the volatility measures were conducted following the same procedure as in Agenor et al., (2004) and Saxegaard (2006).

In the previous chapter, when examining the determinants of excess liquidity (bank-level model) in the involuntary equation, the Strength of Legal Rights Index (LRI) was used to proxy institutional and market conditions in the transition context. Again, balanced data are required to implement a SUR technique and, since this index started to be published relatively late, we thus refrained from using the same proxy in the macro-level model. Since, institutional factors are one of the most important factors in explaining the involuntary excess liquidity, another similar World Bank measure was used, the Rule of Law Estimate (RLE). The RLE for Czech and Estonia is available from 2002-2012 (or earlier), whereas for Kosovo from 2003-2012. However, even though this estimate has a longer span than the RLI, this measure still lacks data for 2013 for all three countries in the sample. Furthermore, the RLE is estimated annually, whereas our dataset is of quarterly frequency. Since we could not find a quarterly institutional proxy from the World Bank or other sources, to be able to use the RLE measure, we have transformed it from an annual to a quarterly indicator as follows: a) initially, each annual value was assumed to be estimated in the fourth quarter of each year, thus each annual value was used as the fourth quarter value and the first three quarters were imputed with cubic spline interpolation approach and b) the last three quarters for the missing year 2013 were also imputed with the cubic spline interpolation method. This approach in the case of the RLE variable may be justifiable since it represents an institutional factor that typically moves in a steady and predictable manner yearly and more so in quarterly frequency. For example, for Czech, Estonia and Kosovo, the actual data at the end of the series in 2012Q4 end with the respective values: 1.006, 1.159 and -0.563, whereas the imputed data for 2013Q1 start with the respective values of 0.989, 1.127 and -0.589 and end in 2013Q4 with the imputed values of 1.002, 1.156 and -0.569, respectively. As it can be seen, the changes in actual values of the data as well those in the imputed ones are relatively small and predictable.

A similar procedure was followed regarding the Eurozone output gap variable. A publicly available output gap for the Eurozone is available in the OECD dataset, however the estimate is provided only in annual frequency. Thus, differently from the RLE estimate where no other data were available for the missing year, in the case of the OECD output gap there are actual data for one quarter of each year. Hence, to transform the annual output gap into a quarterly one, the annual values of the output gap were used in the fourth quarter, whereas the first three quarters were interpolated with the cubic spline interpolation approach. In addition to the OECD output gap for Eurozone in combination with the interpolated values, for a robustness check we also generated our own measure of the output gap for the Eurozone, using the Hodrick-Prescott filter.[106] The filtering was conducted using a univarite approach. The GDP of the Eurozone, as explained in chapter 3, was initially seasonally adjusted, as is commonly done in the business cycle literature, given that seasonal adjustment helps in generating better turning points in the cycle and reduces serial correlation.

Since the dataset is in quarterly frequency, some of the other variables in the model that usually exhibit a seasonal component in the sample were initially seasonally adjusted. Firstly, seasonal adjustment was carried out because of the quarterly data in the model and, secondly, when the output gap was estimated in chapter 3, the GDP series for the three countries were initially seasonally adjusted and then entered in the filtering process. This means that the output gap of the Czech Republic, Estonia and Kosovo are already seasonally adjusted and in this respect is comparable with other variables that also exhibit seasonal components. The data were seasonally adjusted using the census X-12 additive method (in E-Views 8), which is the most commonly used method in the literature. Another approach to seasonally adjust the data is to include quarterly dummies in each equation to control for seasonality; however this method will substantially reduce the degrees of freedom. Given that our dataset is relatively small and the number of parameters to be estimated relatively large, we refrained from using seasonal adjustment with seasonal dummies.

The other variables that were seasonally adjusted are: EL_t, $CreditF_t$, DEP_t, FDI_t and $Govexp_t$. The reason for seasonally adjusting excess liquidity is because the liquid assets of the banks usually increase during the summer holidays, Christmas and other periods of the year when people usually spend more and thus more transaction payments are realised. The same logic for seasonal adjustment applies for deposits and credit flows, as the base of both usually increases in the second and fourth quarter due to seasonal factors. The FDIs are also prone to seasonal factors such as tourism, holidays or when enterprise activity is higher. Other variables in the dataset were not a priori seasonally adjusted, as they either do not exert seasonal pattern or it makes no sense to seasonally adjust them. Such variables in our model are regulatory, risk measures and institutional variables. Thus, the variables that were not

[106] The Hodrick-Prescott filter is one of the well-established and most widely used filtering methods throughout the business cycle literature. The Kalman filter is considered to be a much powerful filter (especially when using transition economies data). However, a more important added value of using Kalman filter in this research was when we generated our variables of interest, such as the output gap for the Czech Republic, Estonia and especially Kosovo. On the other hand, the output gap of Eurozone in the output gap equation represents only one of the control variables, and as using Kalman filter is commonly known to be time consuming, we have instead used the Hodrick-Prescott filter.

seasonally adjusted are: RRR_t, Eq_t, $VOlD_t$, $VOlC_t$, $VOlgdp_t$, NPL_t and the RLE_t. For example, the RRR_t is a regulatory ratio set by the central bank and usually does not change for many years in a row as is the case with Czech and Kosovo; the Eq_t also mainly represent the regulatory requirements for capital holdings (plus additional equity that banks may want to hold); $VOlD_t$, $VOlC_t$, $VOlgdp_t$ represent the risk perceptions of banks for their customers or the entire economy of which banks are aware and continually monitor regardless of the pronounced seasons or not; the NPL_t represent the non-repaid loans for six months or more, thus measuring the bad debt accumulated in previous periods.[107]

The summary statistics of each variable as described in tables 5.2, 5.3 and 5.4 presented below:

[107] The three countries in the sample use the IMF's definition for loan classification. The loan classification based on the days on the delaying loan repayment according to the IMF's definition is: Standard loans (0 - 30 days), Watch (31 – 60 days), Sub-Standard (61 – 90 days), Doubtful (91 – 180 days) and Loss (< 181 days). The doubtful and loss category comprise the NPLs. The NPLs will comprise the delayed loans for more than 180 days or even delayed for years, therefore they do not particularly depend on the current quarter with the pronounced seasonality.

Table 5.2. Summary statistics, Czech Republic

Country	Variable	Mean	Median	Maximum	Minimum	Standard Deviation	Observations (without interpolated data)
Czech Republic	Excess Liquidity	15.47	14.48	25.66	8.06	4.42	48
	Required Reserve Ratio	2.00	2.00	2.00	2.00	0.00	48
	Deposit Interest Rate	1.20	1.15	2.46	0.63	0.41	48
	Real Growth	-0.02	-0.45	11.82	-5.00	2.82	48
	Volatility of Private Sector Credit	0.25	0.21	1.06	0.00	0.23	48
	Volatility of the Deposits	0.14	0.14	0.33	0.02	0.07	48
	Volatility of GDP	1.00	0.20	6.12	0.01	1.86	48
	Non-Performing Loans	5.47	5.60	11.90	2.60	2.07	48
	Equity	10.56	10.66	11.77	8.87	0.78	48
	Deposits	0.55	0.87	3.43	-3.14	1.24	48
	Government Deposits	0.55	0.59	6.49	-4.67	2.48	48
	Credit	16.02	13.71	138.89	-58.53	31.63	48
	Loan interest rate	5.62	5.62	6.63	4.43	0.51	48
	Rule of Law Estimate	0.25	0.36	3.85	-3.17	2.02	44
	Foreign Direct Investment	4.20	3.91	24.14	-2.09	4.12	48
	Output Gap	0.58	0.67	3.53	-2.24	1.74	48
	a) Eurozone Output Gap - OECD	0.25	0.36	3.85	-3.17	2.02	48
	b) Eurozone Output Gap - HP filter	-0.01	-0.11	3.65	-3.72	1.57	48
	Government Expenditure	42.74	42.60	58.90	36.20	3.63	48

Table 5.3. Summary statistics, Estonia

Country	Variable	Mean	Median	Maximum	Minimum	Standard Deviation	Observations (without interpolated data)
Estonia	Excess Liquidity	5.82	2.87	25.66	-6.53	9.70	48
	Required Reserve Ratio	10.98	13.00	15.00	2.00	5.13	48
	Deposit Interest Rate	2.81	3.00	4.52	0.41	1.19	48
	Real Growth	1.17	0.74	7.86	-2.16	2.13	48
	Volatility of Private Sector Credit	0.21	0.08	0.71	0.00	0.24	48
	Volatility of the Deposits	0.99	0.85	3.18	0.10	0.74	48
	Volatility of GDP	0.10	0.08	0.30	0.01	0.07	48
	Non-Performing Loans	2.77	2.02	7.40	0.30	2.37	48
	Equity	6.02	5.78	12.74	1.51	3.95	48
	Deposits	6.02	5.35	22.30	-4.91	5.24	48
	Government Deposits	0.55	0.59	6.49	-4.67	2.48	48
	Credit	7.56	5.04	29.43	-6.81	9.44	48
	Loan interest rate	7.19	7.18	13.27	3.94	2.47	48
	Rule of Law Estimate	1.00	1.09	1.16	0.73	0.15	44
	Foreign Direct Investment	4.20	3.91	24.14	-2.09	4.12	48
	Output Gap	0.78	3.20	6.19	-5.53	4.17	48
	a) Eurozone Output Gap - OECD	0.25	0.36	3.85	-3.17	2.02	48
	b) Eurozone Output Gap - HP filter	-0.01	-0.11	3.65	-3.72	1.57	48
	Government Expenditure	36.74	36.60	46.83	30.26	4.29	48

Table 5.4. Summary statistics, Kosovo

Country	Variable	Mean	Median	Maximum	Minimum	Standard Deviation	Observations (without interpolated data)
Kosovo	Excess Liquidity	28.28	21.05	88.62	13.33	18.29	40
	Required Reserve Ratio	10.00	10.00	10.00	10.00	0.00	40
	Deposit Interest Rate	3.48	3.48	4.65	2.68	0.53	39
	Real Growth	0.92	0.87	2.35	-0.26	0.52	40
	Volatility of Private Sector Credit	0.96	0.65	3.45	0.07	0.79	40
	Volatility of the Deposits	0.96	0.63	5.02	0.07	1.30	40
	Volatility of GDP	1.06	0.99	2.67	0.01	0.67	40
	Non-Performing Loans	4.95	4.60	8.50	1.70	1.79	32
	Equity	8.73	9.08	10.63	5.64	1.17	40
	Deposits	3.86	4.07	14.68	-21.17	5.13	48
	Government Deposits	-0.03	-0.01	14.15	-6.15	2.66	48
	Credit	3.79	3.75	8.68	-0.17	2.10	40
	Loan interest rate	14.27	14.57	15.52	11.66	0.91	39
	Rule of Law Estimate	-0.79	-0.76	-0.55	-1.18	0.20	36*
	Foreign Direct Investment	8.84	8.61	18.28	2.78	3.53	28*
	Output Gap	-1.90	-2.49	3.89	-5.79	2.69	40
	a) Eurozone Output Gap - OECD	-0.50	0.08	3.27	-3.99	2.34	40
	b) Eurozone Output Gap - HP filter	-0.01	-0.11	3.65	-3.72	1.57	40
	Government Expenditure	48.67	47.03	65.91	25.45	11.61	40

*Note: The complete period of the series in the sample is available in yearly frequency, but some of the quarters are interpolated.

It should be noted that two other variables were considered to be included in the involuntary excess liquidity equation, namely donor funds and the remittances from abroad, given their particular importance in the Kosovo economy. However, the donor funds mainly circulate via the Treasury at the central bank, meaning that they do not necessarily contribute in excess liquidity build up. The total remittances, on the other side, comprise a considerable portion of GDP, but nonetheless only about one third circulate via the banking system. Furthermore, the remittance funds in the banks are already counted in total deposits, thus to avoid double-counting, we refrained from exploring their particular effect on the involuntary excess liquidity. Thus, after due consideration, donor funds and remittances were not included in the model.

Finally, even though it may be an issue largely neglected in published work, it should be acknowledged that our estimates may suffer from measurement error. The literature usually acknowledges two types of measurement errors: the one coming from the technical statistical issues and the second from a theoretical and conceptual misspecification. Given the issue of the missing values and of the interpolation and imputation in our dataset, our estimates are more likely to suffer from statistical inaccuracies. “But no economic model is perfectly true, and no economic data are perfectly measured" (Cochrane, 2012, p. 5). One way to address the measurement error is to instrument the variables with potential measurement error. However, given the availability of our data, especially for Kosovo, instrumenting is not a feasible strategy. However, to add confidence to the derived conclusions, a multiple robustness checks will be performed, with and without interpolated data. Three sets of the SUR estimations with a different time span will be carried out: a) the SUR estimation using three countries with the complete time-series (2004Q1-2013Q4) combining actual and interpolated data; b) the SUR estimation with actual data using a relatively smaller dataset with a shorter time-series (2002Q1-2012Q4) and with Czech and Estonia only, since these two countries do not have missing values in the series, except for the rule of law estimate and c) all three countries in the sample, with the actual data and without the interpolated data but with a shorter time span (2007Q1-2012Q4).

In the following section the unit root tests and the estimated results will be discussed.

5.5. Unit root tests for the SUR model

Given the distinction between stationary, trend-stationary and unit root processes in time-series data, it is important to distinguish if a particular series generally increasing in value is being driven by some underlying trend or whether its evolution reflects a unit root in its data generating process. The unit root tests in our case are not only important because time-series data are being used, but also, as will be discussed in section 5.6.2, the SUR modelling requires series to be stationary. As usual, when testing each series for the presence of a unit root, different kind of tests generate inconclusive results. Therefore, in the following a few types of unit tests were carried out. All of the unit root tests assume white noise errors in the regression, i.e. no serial correlation in the residuals, otherwise the tests will lose significant power in rejecting the null of a unit root (Wooldridge, 2002, p.581). To cope with this, each specific series in the model was tested with the Breusch-Godfrey LM test with the corresponding number of lags to remove serial correlation (the last column in the tables below), then the series were tested for unit roots.[108]

Initially, each series was tested using the most commonly used test, the Augmented Dickey-Fuller (ADF) test with a constant or drift term. In addition, the series were also tested by the

[108] The full output of the unit root tests is available from the author upon request.

DF-GLS approach, which though similar to the ADF test is considered to be more powerful as it firstly transforms the series via a generalized least square (GLS) regression before performing the test (StataCorp, 13). Also, differently from the ADF test, the DF-GLS test is performed on GLS-detrended data, thus controlling for whether the series has unit a root as opposed to the possibility of being trend-stationary. However, while testing for the presence of a unit root, none of these tests control for possible structural breaks in the series. Testing for possible breaks in unit root testing is particularly important since in their presence the ADF and DF-GLS (or other similar tests) may wrongly lead to the conclusion that the series has a unit root. This is because "the model excluding structural change is clearly misspecified by the omission of relevant explanatory variables" (Blackwell, 2005, p. 56). For example, a series that undergoes an intercept shift is not covariance stationary, but it can be made so if regressed with an intercept shift dummy that identifies the shift period (Wooldridge, 2002, p. 348). In the case of transition economies more than one structural break may be present. However, the main model dataset starts from 2004, which represents a period when most of the transition economies had already accomplished their main structural reforms in the 1990s. Thus, for the period being examined an obvious structural change that we need to account for in all the series is the period right after the onset of the financial crisis. Assuming that the break period (T_b) is known, the Perron unit root test was conducted for each series manually.[109] To account for all three possible types of breaks, each series was tested in the presence of a one-time break dummy (1 in the T_b, 0 otherwise), an intercept shift dummy (0 up to the T_b, 1 from T_b+1 onwards) and a trend shift dummy (0 up to the T_b, from T_b+1 a deterministic trend). Even though the onset of the financial crisis is generally agreed to have happened around 2008, different countries experienced its impact earlier or with a lagged effect. This is why the break period for the three countries in our sample is slightly different (Czech 2008Q3, Estonia 2008Q1 and Kosovo 2009Q1). In the Czech Republic case, which still follows a conservative banking model mainly relying on domestic funds, economic stagnation started in 2008. Being an export-led country, the decline in foreign demand, especially from Germany, started to show its impact in the second half of 2008. However, only in the third quarter of 2008 did the Czech government start to form fiscal rescue plans. In Estonia, the severe effects of the financial crisis started to show at the end 2007 and early 2008 mainly due to the real estate bubble building up during the preceding years, as well as the slump in domestic consumption and investment. Kosovo on the other side was growing

[109] The full output of the Perron unit root tests is available from the author upon request.

steadily up to 2008. As a country with a simple banking model which also relied mainly on domestic funds and relatively low external exposure, it was one of the few countries in Europe not to experience recession as a consequence of the financial crisis. However, as a small open economy, the effects of the financial crisis started to show with a lag effect at the beginning of 2009 mainly via the external sector, such as the decline in FDI inflows, remittances and exports. However, one of the weaknesses of the Perron unit root test is its inability to deal with more than one break.[110] In the case of transition economies, there may be more than one structural break in the series. Thus, to extend our unit root tests with structural breaks, we have also implemented the Clemente et al., (1998) unit root test with two possible breaks within the history of each series using the c*lemio2* testing program in Stata 12. The *clemio2* program controls for two one-time change breaks and two gradual shifts in the mean of the series. There were cases where none of the structural breaks turned out as significant, or other cases when both of the identified breaks from the programme were significant.

Given that the results generated from different tests for unit roots were inconclusive in many cases, the series were considered to be stationary – I(0), or unit root –I(1), mainly depending on the tests with structural breaks. However, when the two unit root tests with structural breaks also generated inconclusive results or contradictory results then the series were also judged based on the visual inspection of the series. Thus, mainly based on the unit root test results, the following variables were considered to be I(0):

Czech: *RRR, iD, VolGDP, NPL, growth, FDI, OG, OGEU, Govexp*
Estonia: *RRR, VolDep, growth, Dep, FDI, OG, OGEU*
Kosovo: *RRR, VolDep, growth, Dep, Govdep, OG, OGEU*

Whereas the rest of the variables are treated as I(1) thus differenced for estimation and interpreted as changes in the respective ratios.

---

[110] In the absence of an automated procedure, the Perron test was manually implemented, given that the time of the break is known. However, there are automated procedures to test for one break in the series, such as the Zivot and Andrews (1992) test (*zandrews*) and the Clemente et al., (1998) test via the programs: *clemio1* (intercept shift break) or *clemao1* (one-time break dummy); the similar Clemente's test with two breaks is implemented via the *clemio2* and *clemao2* programs in Stata.

The tested results for each series of the three countries are presented in Tables 5.5, 5.6 and 5.7 below.

Table 5.5. Unit root tests, Czech Republic

	Variable symbol	DF-GLS+ Trend	ADF + Constant	ADF + Drift	Decision 1	Perron with structural breaks tests	Decision 2	Clemio 2 structural breaks	Decision 3	Lags for ser. Corr.
Czech	*EL*	Not rejected	Not rejected	Rejected	Probably non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	2
	ΔEL	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	2
	*RRR**	-	-	-	-	-	-	-	-	-
	rgrowth	Rejected	Rejected	Rejected	Stationary	Rejected	Stationary	Rejected	Stationary	1
	i^D	Not rejected	Not rejected	Rejected	Probably non-stationary	Rejected	Stationary	Rejected	Stationary	2
	$\Delta\, i^D$	Rejected	Rejected	-	Stationary	-	-	-	-	2
	VolC	Not rejected	Not rejected	Not rejected	Non-stationary	Rejected	Stationary	Not rejected	Non-stationary	1
	Δ VolC	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	VolD	Not rejected	Not rejected	Rejected	bably non-statio	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ VolD	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	VolGDP	Rejected	Rejected	Rejected	Stationary	Rejected	Stationary	Rejected	Stationary	0
	NPL	Rejected	Rejected	Rejected	Stationary	Not rejected	Non-stationary	Rejected	Stationary	1
	Eq	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ Eq	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	Dep	Rejected	Rejected	Rejected	Stationary	Rejected	Stationary	Not rejected	Non-stationary	0
	Δ Dep	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	0
	Govdep	Not rejected	Rejected	Not rejected	Probably non-stationary	Not rejected	Non-stationary	Rejected	Stationary	1
	Δ Govdep	Not rejected	Rejected	-	Possibly stationary	-	-	-	-	1
	Cred	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ Cred	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	i^L	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	2
	$\Delta\, i^L$	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	2
	RLE	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	2
	Δ RLE	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	2
	FDI	Rejected	Rejected	Rejected	Stationary	Not rejected	Non-stationary	Rejected	Stationary	0
	OG	Rejected	Rejected	Rejected	Stationary	-	-	Rejected	Stationary	2
	Govexp	Rejected	Rejected	Rejected	Stationary	Not rejected	Non-stationary	Rejected	Stationary	0
	*OGEU (OECD)***	Rejected	Rejected	Rejected	Stationary	-	-	-	-	5
	*OGEU (HP)***	Rejected	Rejected	Rejected	Stationary	-	-	-	-	1

Note: *Even though prior to 2002 the RRR in Czech Republic was higher and then started to slowly to redue to converge with the ECB regulation of 2% on average, from the beginning of the series 2002 to 2013, the RRR in Czech stood unchanged at 2% throughout the period, thus the unit root tests did not perform in Stata 12.
**The Output Gap estimates of the Eurozone are the same ones used for all three countries, thus its tests are only presented in this table.

Table 5.6. Unit root tests, Estonia

	Variable symbol	DF-GLS+ Trend	ADF + Constant	ADF + Drift	Decision 1	Perron with structural brekas tests	Decision 2	Clemio 2 structural breaks	Decision 3	Lags for ser. Corr.
Estonia	*EL*	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	ΔEL	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	*RRR**	Not rejected	Not rejected	Not rejected	Non-stationary	-	-	Rejected	Stationary	2
	Δ RRR	Rejected	Rejected	-	Stationary	-	-	-	-	2
	rgrowth	Rejected	Rejected	Rejected	Stationary	Rejected	Stationary	Rejected	Stationary	1
	i^D	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Rejected	Stationary	1
	Δi^D	Rejected	Rejected	-	Stationary	-	-	-	-	1
	VolC	Not rejected	Not rejected	Rejected	Probably non-stationary	Rejected	Stationary	Not rejected	Non-stationary	3
	Δ VolC	Rejected	Rejected		Stationary	-	-	Rejected	Stationary	3
	VolD	Not rejected	Not rejected	Rejected	Probably non-stationary	Rejected	Stationary	Rejected	Stationary	3
	Δ VolD	Not rejected	Rejected	-	Probably non-stationary	-	-	-	-	3
	ΔΔ VolD	Rejected	-	-	Stationary	-	-	-	-	3
	VolGDP	Not rejected	Not rejected	Rejected	Probably non-stationary	t = CV	Probably non-stationary	Not rejected	Non-stationary	1
	Δ VolGDP	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	NPL	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	2
	Δ NPL	Not rejected	Not rejected		Non-stationary	-	-	Rejected	Stationary	2
	ΔΔ NPL	Rejected	Rejected		Stationary	-	-	-	-	2
	Eq	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Rejected	Stationary	1
	Δ Eq	Rejected	Rejected		Stationary	-	-	-	-	1
	Dep	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Rejected	Stationary	2
	Δ Dep	Rejected	Rejected		Stationary	-	-	-	-	2
	Govep	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ Govdep	Not rejected	Rejected	-	Possibly stationary	-	-	Rejected	Stationary	1
	Cred	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	2
	Δ Cred	Rejected	Rejected		Stationary	-	-	Rejected	Stationary	2
	i^L	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δi^L	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	RLE	Rejected	Not rejected	Rejected	ossibly stationar	Not rejected	Non-stationary	Not rejected	Non-stationary	4
	Δ RLE	-	Rejected	-		-	-	Rejected	Stationary	4
	FDI	Rejected	Rejected	Rejected	Stationary	Rejected	Stationary	Not rejected	Non-stationary	0
	OG	Rejected	Rejected	Rejected	-	-	-	Rejected	Stationary	1
	Govexp	Not rejected	Not rejected	Rejected	Probably non-stationary	Not rejected	Non-stationary	Rejected	Stationary	2
	Δ Govexp	Rejected	Rejected	-	Stationary	-	-	-	-	2

*The RRR in Estonia has two obvious structural changes (policy induced) therefore only the Clemio unit root test with two structural breaks was performed.

Table 5.7. Unit root tests, Kosovo

	Variable symbol	DF-GLS+ Trend	ADF + Constant	ADF + Drift	Decision 1	Perron with structural brekas tests	Decision 2	Clemio 2 structural breaks	Decision 3	Lags for ser. Corr.
Kosovo	*EL*	Not rejected	Rejected	Rejected	Possibly stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	3
	ΔEL	Rejected	-	-	-	-	-	Rejected	Stationary	3
	*RRR**	-	-	-	-	-	-	-	-	-
	rgrowth	Rejected	Rejected	Rejected	Stationary	Not rejected	Non-stationary	Rejected	Stationary	1
	i^D	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δi^D	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	VolC	Not rejected	Not rejected	Rejected	Probably non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ VolC	Rejected	Rejected		Stationary	-	-	Rejected	Stationary	1
	VolD	Not rejected	Rejected	Rejected	Possibly stationary	Not rejected	Non-stationary	Rejected	Stationary	1
	Δ VolD	Rejected	-	-	Stationary	-	-	-	-	1
	VolGDP	Not rejected	Not rejected	Rejected	Probably non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ VolGDP	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	NPL	Rejected	Not rejected	Not rejected	Probably non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ NPL	-	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	Eq	Not rejected	Not rejected	Rejected	Probably non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ Eq	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	Dep	Rejected	Rejected	Rejected	Stationary	Not rejected	Non-stationary	Rejected	Stationary	0
	Govdep	Rejected	Rejected	Rejected	Stationary	Not rejected	Non-stationary	Rejected	Stationary	1
	Cred	Not rejected	Not rejected	Rejected	Probably non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ Cred	Rejected	Rejected	-	Stationary	-	-	Rejected	Stationary	1
	i^L	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	2
	Δi^L	Rejected	Rejected		Stationary	-	-	Rejected	Stationary	2
	RLE	Rejected	Not rejected	Rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	8
	Δ RLE	Rejected	Not rejected	-	Probably non-stationary	-	-	Not rejected	Non-stationary	8
	ΔΔ RLE		Rejected	-	Stationary	-	-	Rejected	Stationary	8
	FDI	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Rejected	Stationary	1
	Δ FDI	Rejected	Rejected		Stationary	-	-	-	-	1
	OG	Rejected	Rejected	Rejected	Stationary	-	-	Rejected	Stationary	2
	Govexp	Not rejected	Not rejected	Not rejected	Non-stationary	Not rejected	Non-stationary	Not rejected	Non-stationary	1
	Δ Govexp	Rejected	Rejected		Stationary	-	-	Rejected	Stationary	1

*The RRR in Kosovo is 10% throughout the period, thus the unit root tests did not perform in Stata 12.

5.6. Methodology

5.6.1. Single equation approach: FGLS and PCSE

Given that our model consists of three countries, namely Czech Republic, Estonia and Kosovo, initially the dynamic linear regression model was specified using time-series data for each country. Since, initial hypothesis was constructed on the basis that excess liquidity was expected to have an impact on the output gap, the output gap was set the dependent variable and excess liquidity as the main explanatory variable, in addition to other control variables (year dummies, interaction terms etc.). However, the time-series models were misspecified, mainly due to serial correlation and omitted variables. Further on, a linear regression model and a fixed effects model within a panel framework was specified, however as it turned out, the model suffered from heteroscedasticity, serial correlation and a relatively high cross-country correlation of the error terms, thus these model specifications yielded misspecified estimates.

In an attempt to find an appropriate method that deals with error structures, the panel model was initially estimated using the so called Feasible Generalised Least Square (FGLS) model originally proposed by the Parks (1967). The FGLS model basically represents a transformed form of the OLS; the GLS transforms the linear regression model with a general error covariance matrix Ω to another linear regression model where the error covariance matrix would be suitable for OLS estimation, i.e. so that Ω becomes 'scalar' (Ooms, 2006/2007, p. 9 – 22).[111] Furthermore, the FGLS approach initially seemed appropriate for estimating the panel model, given that it was widely used in small datasets (especially in political economy studies) and requires the number of T periods to be much larger than the number of cross-units N, which is our case ($N = 3$, $T = 50$ per equation of each country), because all the pairwise contemporaneous covariances are estimated by (N(N+1)/2, denoted by the matrix $\widehat{\sum}$ using NT observations (N – nr. of units; T – nr. of periods). $N(N+1)/2 = 6$ parameters of the covariance matrix to be estimated, where each parameter is estimated using $2T/N = 33$ observations, meaning that our model fulfils the criteria of $T >> N$ and has a relatively large number of observations available for estimation of each covariance matrix parameter.

Later, Becks and Katz (1995) strongly criticize the use of the FGLS approach for generating overconfident results and three times lower standard errors. As an alternative to the rather 'complex' FGLS approach, Becks and Katz (1995) propose another, simpler method, Panel Corrected Standard Errors (PCSE) estimation, which is constructed to fit linear cross-sectional time-series models when the errors are *not* assumed to be independent and identically distributed (*i.i.d*),. Unlike FGLS, the PCSE approach does not require many assumptions to be imposed in the model. Beck and Katz (1995) argue that a superior way to handle the complex error structures in TSCS analysis is to estimate the coefficients by OLS and then compute PCSEs. In this approach, Ω is a $NT \times NT$ block diagonal matrix with $\sum$, an $N x N$ matrix of contemporaneous correlations across the diagonal. OLS residuals, denoted $e_{i,t}$ for unit i at time t, are used to estimate the elements of Σ; then the standard errors of the

[111] To envisage the step-by-step transformation of the estimates and the error terms in the GLS framework, see Woodridge (2002, pp. 263 – 267) or to manually implement the GLS approach see p. 268. The lecture of Winship (2009) provides a general and intuitive understanding of the GLS in general.

coefficients are computed using square roots of the diagonal elements. Although this approach estimates the same number of parameters as the FGLS method, the authors argue that it has better small sample properties.

Overall, given that the two approaches, namely FGLS and the PCSE share common characteristics, such as: both estimators are considered to be consistent, if the conditional mean ($\beta Xi,j$) is correctly specified; both estimate the $\widehat{\sum}$ from the OLS estimates in the first stage, both estimate same number of parameters, both require balanced dataset to parameterize the covariance matrices etc., the panel model was estimated with both techniques in parallel.

Given the complex error structure, the FGLS and PCSE initially appeared to be the appropriate models available to estimate our model. The estimated results from both techniques generated quantitatively different results but qualitatively fairly similar ones. Further, the standard errors using the FGLS method were not three times smaller than the OLS as claimed by Becks and Katz (1995), thus the Becks and Katz finding appear not be universal. As a result researchers should report both models.

As the initial intention was to uncover a causal relationship between output gap and excess liquidity, initially the output gap was set as the dependent variable whereas excess liquidity as the main explanatory variable. Furthermore, suspecting possible simultaneity between the output gap and excess liquidity, we have also tried putting excess liquidity as a dependent variable, whereas the output gap as the main explanatory variables.

Whereas, the estimated results generated these findings:

a) in current values, output gap (independent variable) determines excess liquidity (dependent variable) as strongly as excess liquidity (independent variable) determines output gap (dependent variable); and

b) the second lag (and other) lagged values of the output gap determine excess liquidity, while second lagged values of excess liquidity determine the output gap.

A selected part of the FGLS and PCSE estimates is reported in Appendix 5.1.

Therefore, the results confirmed perfect simultaneity between output gap and excess liquidity:

To conclude, for the particular relationship between output gap and excess liquidity, the theoretical reasoning together with our empirical confirm that pursuing a single equation approach where one variable (output gap) is a dependent variable, whereas the other (excess liquidity) is independent variable is a conceptually misconceived. This is because, no matter which single-equation technique was applied, thinking that it would be appropriate to model the characteristics of our dataset, the results would always seem to indicate a correlation relationship. Further, even though the initial intention was establishing a causal relationship causal or bi-direction causation between output gap and excess liquidity, in a correlation bound relationship it would still not be possible to distinguish which the variable is causing the other first. Therefore, both theory (sections 5.2. and 5.3) and empirical evidence (section 5.6.2) confirm the need for another modelling technique that takes into account a series of variables that are endogenous and correlated in nature, i.e. a system-equation approach. Following these suggestions, we will switch into another modelling technique, that of seemingly unrelated regressions (SUR).

5.6.2. System equation approach: SUR

The prevailing theory on growth and finance, as discussed in sections 5.2 and 5.3, proclaims a causal relationship from finance to growth or, in a few cases, a simultaneous relationship. Most of the studies supporting this view are represented by prominent authors in prestigious journals, and – to the author's knowledge – has not yet been challenged. While seeking to define the relationship between the output gap and excess liquidity we, too, began our investigation by hypothesising a causal relationship. Even Agenor et al. (2004) uses the output gap as an independent variable in their model of precautionary excess liquidity. However, because they were unable to explain the counterintuitive coefficient sign obtained in estimation, they dropped the output gap from their model of excess liquidity. Indeed, Agenor et al. (2004) could not find a causal relationship between the two. The experience of Agenor et al. (2004), together with our own inability to identify a directly causal relationship between excess liquidity and the output gap, led us first to theoretical reconsideration of the relationship between these two variables and then to adopt a correspondingly different empirical approach. Even though we make no claim that our arguments on their relationship are definitive, the position that we arrive at provides a new direction for thinking about excess liquidity and the output gap.

The presence of a (negative) output gap indicates that the economy could produce more with the existing resources, i.e. that there are underutilised resources. In addition, excess liquidity means that there are underutilised savings in the banking system. Hence, as discussed in section 5.3, they both represent different indicators of underutilised resources; or two different perspectives that describe the general state of the economy, which in a Keynesian model reflects equilibrium in a depressed economy. This is why, conceptually, the output gap and excess liquidity cannot cause each other but, rather, are both potentially endogenous variables caused by a similar range of underlying determinants. In this case, excess liquidity and the output gap are correlated, because both are jointly determined within the same wider system; but they do not determine one another. The causation comes from that deeper underlying system, as presented in figure 5.3 below, of common observed and unobserved determinants. Because of common determinants, the output gap and excess liquidity are both endogenous variables, not because of simultaneity but because of joint determination within a wider system. In this system, excess liquidity does not belong in the equation of the output gap, just as the output gap does not belong on the right-hand-side of the equation for excess liquidity. Instead, both represent dependent variables within a wider system. Therefore, as a corollary of this theoretical reconsideration, we examine this relationship with its empirical counterpart, that of Seemingly Unrelated Regressions (SUR).

As discussed in the previous section, in this section we have moved away from relying on estimates from a single-equation approach to specify a systemic model. To take account of the correlated error terms of the equations with potentially endogenous dependent variables, this research employs a different and arguably more appropriate estimation method, that of Zellner's (1962) system-equation approach, the Seemingly Unrelated Regressions (SUR). This technique has not been previously used in our field of research in Kosovo's and is not encountered in the empirical studies analysing similar topics, even for the developed economies. After the assessment of the applicability of the various estimation methods discussed in the previous section, given the specific nature of the data series (correlation) and the possible phenomenon of contemporaneous correlation among the countries, we have selected Zellner's (1962 and 1963) Seemingly Unrelated Regression (SUR) model.

The general form of the SUR model can be presented with the following system of equations:

$$Y_{1t} = \beta'_1 x_{1t} + u_{1t}$$
$$Y_{2t} = \beta'_2 x_{2t} + u_{2t}$$
$$\vdots$$
$$Y_{nt} = \beta'_n x_{nt} + u_{nt} \qquad (5.1)$$

Where: Y is the dependent variable; β' is a vector of coefficients; X is a matrix of independent variables; u is the error terms; and n and t are cross-sectional and time specific subscripts.

The above equations can be stacked as a large equation system and can be presented more compactly as follows:

$$Y_t = X_t'\beta + u_t \qquad (5.2)$$

Where: Y_t is $TN \ x \ 1$ matrix of dependent variables; X' is a $TN \ x \ K$ matrix of independent variables; β is a $K \ x \ 1$ matrix of coefficients; u is $TN \ x \ 1$ matrix of error terms; T and N are the number of time and cross-sectional observations, respectively and K is the number of independent variables.

The rationale for selecting this model is based upon several reasons. Firstly, in the case when there is contemporaneous correlation among the disturbances that are by nature heteroskedastic, then the SUR model based on a Feasible Generalised Least Squares (FGLS) estimator (see section 5.6.1 for a description on the FGLS properties) provides more efficient estimates compared to OLS, by using the information of the variance-covariance matrix of the error terms (Becks and Katz, 1995). This is not done to deal with potential bias, because a nonzero error covariance does not cause bias in the parameter estimates; rather, it causes a loss of efficiency (Cameron, 2005, p. 353). Thus, when the correlation among the error terms of each cross-sectional unit is higher, then the FGLS estimator is able to use more information from the variance-covariance matrix of the error terms and hence, the efficiency gain by employing the SUR model will be higher (Vogelvang, 2005; Baum, 2006 and Greene 2008). There is no qualitative difference between OLS and SUR FGLS, but the latter gives much lower standard errors (Bhattarai, 2011). However, with identical regressors in each equation, the weight variation disappears and the SUR estimator is identical to OLS. Secondly, as

Maddala (1997) and Moon and Perron (2005) argue, it is designed for samples with a large time dimension (T) and small or finite cross sectional dimension (N) where one of the major requirements is T to be substantially greater than N, which is the case with our data (T = 38 per equation; N=3)[112]. Thirdly, it can estimate different slope coefficients for each cross-sectional unit, which allows for their cross-sectional equality to be tested. This then enables us to investigate whether there is heterogeneity among countries in our sample regarding the output gap and excess liquidity relationship and what their major determinants are. This option will actually enable us to test whether the slope coefficients statistically differ among the cross-sectional units. Fourthly, another advantage of employing the SUR model is that in the case, when the repeated iterations in calculating the coefficients and their variances for each cross-sectional unit converge, then the FGLS estimator becomes equal to the Maximum Likelihood Estimator (MLE). This may provide some additional efficiency gains, under the condition that the normality assumption about error terms is fulfilled (Greene, 2008 and Moon and Perron, 2006). However, as discussed in Greene (2008, p. 347), whether MLE provides efficiency gains in small samples is uncertain and depends on the specificity of the dataset.

The SUR model has some limitations and requires certain assumptions to be fulfilled. The main assumptions are the exogeneity of the regressors and a normal distribution of the residuals; the latter assumption is mainly for MLE but not for FGLS. As Zellner (1963, p. 988) argues: "….even when normality is not present, the estimation procedure is applicable and will yield consistent coefficient estimators which are asymptotically normally distributed." In respect of the exogeneity assumption, the strongest form is the strict exogeneity assumption where all regressors from each equation are uncorrelated with the respective equation's error terms for all time periods:

$$E = (u_t \mid x_1, x_2, x_3, \ldots\ldots, x_t) = 0 \tag{5.3}$$

However, Wooldridge (2002, p. 318) argues that this assumption may be relaxed by assuming a contemporaneous exogeneity, i.e. no correlation between the regressors and the error terms in the same time period, as presented below:

$$E = (u_t \mid x_t) = 0 \tag{5.4}$$

[112] Zellner (1962) in his original paper about SUR estimation technique uses a sample of 2 firms.

The major limitation of the SUR model is that it does not properly deal with non-stationary variables, because cointegration methods are not developed within the framework, while dynamic specifications are still in the process of development (Wooldridge 2010, p. 201). Another limitation is that if any of the system equations is miss-specified, then all coefficients in each equation will be inconsistently estimated. Therefore, for a consistency check, it is argued that the results should be compared with the ones estimated with the OLS conducted on an equation-by-equation basis (Moon and Perron 2006).

Regarding the issue whether the SUR model provides an appropriate estimator for dynamic models, we have considered if the recent work in the field of dynamic SUR is applicable in our case. More precisely, we have assessed the estimation method applied in Sorensen and Werner (2006) who employ Dynamic SUR (DSUR) for estimating the long-run relationship among the variables and the ECM for SUR (SURECM) for estimating the short-run dynamics. The recently developed DSUR method by Mark et al. (2005) and Moon and Perron (2005) is based on the dynamic ordinary least squares (DOLS) method by Stock and Watson (1993) that controls for the possible endogeneity of the regressors. Namely, "System DOLS is distinguished from ordinary DOLS in that endogeneity in equation *i* is corrected by introducing leads and lags of the first difference not only of the regressors of equation *i* but also of the regressors from all other equations in the system." (Mark et al., 2005, p.798). However, estimating a Dynamic SUR as a system may absorb too many degrees of freedom due to the large number of leads and lags that have to be included in each equation due to the exogeneity requirement of the SUR approach. Even without the lags and leads, we have around 70 parameters to be estimated and only 150 observations, which suggests that this method may not applicable given our dataset.

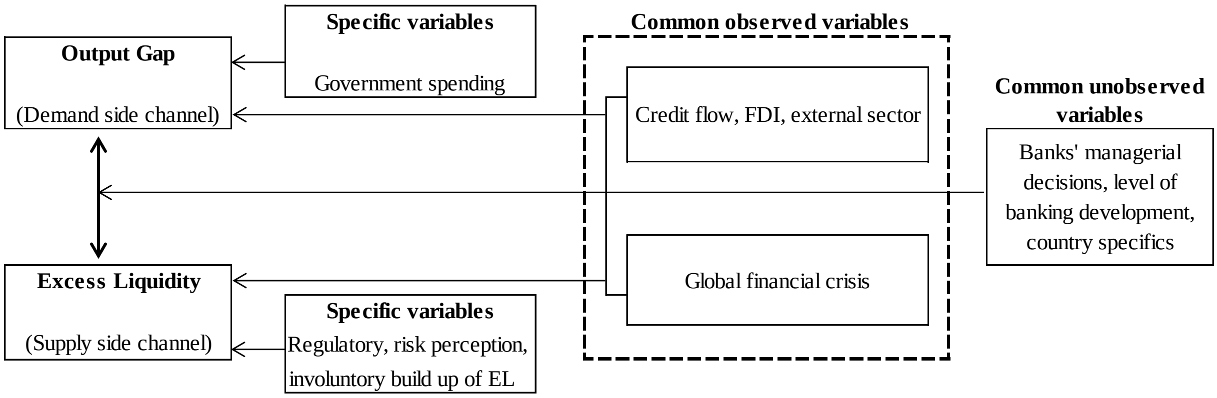
Regarding the SURECM method for estimating the short-run dynamics, we argue that it also has some severe weaknesses and therefore, is not appropriate in our case. This method is based on the Engle-Granger error correction model, but again does not test for the stationarity of the residuals from the long-run relationship equation. The authors that employ this method (Thomson et al., 2002; Kim, 2004 and Sorensen and Werner, 2006), make a simple assumption that the residuals from the long-run relationship equation are stationary and thus, the variables are cointegrated as suggested by economic theory, and proceed by estimating the short-run dynamics of the model.

In the following, the relationship between the output gap and excess liquidity within the context of the SUR is presented.

5.6.3. The SUR in the output gap and excess liquidity framework

We begin the discussion on application of the SUR within our macroeconomic framework between output gap and excess liquidity (only three countries in the sample). As we discussed in section 5.6.1, trying to investigate the impact of excess liquidity on the output gap, initially we treated the output gap as a dependent and the excess liquidity as an independent variable. However, no matter what the methodology used or the choice of the variables, the most that was generated from a single equation approach was a correlation relationship between the two. This is because, as theory suggests (Biggs and Mayer, 2013) and as empirically confirmed (see sections 5.6.3 and 5.7), these two variables are brought together by common underlying determinants, both observed and unobserved. The common observed and manifested unobserved determinants may be correlated between the error terms of the output gap and excess liquidity regression equations, in which case we can gain more efficient estimators by estimating the all the equations representing output gap and excess liquidity jointly, as was shown by Zellner (1962). However, aside from this conceptual relationship, the two variables may have other connections to one another, which would support our proposition of the output gap and excess liquidity being jointly determined within a wider system. They both may be an outcome of a wider system, but not because they are causing each-other. Since they are both potentially endogenous, they may both represent dependent variables within a wider system. Since we are interested in the effect of both variables the SUR method provides a systematic framework for estimation of the determinants of the output gap and excess liquidity. In other words, the system-equation approach via the SUR method enables us to test between mutual exogeneity and mutual endogeneity (Efendic et al., 2011) as the characteristic relationship between the output gap and excess liquidity, as presented in figure 5.3.

Figure 5.3. Visual depiction of the SUR



Source: Author's illustration.

As depicted in figure 5.3, the output gap and excess liquidity may appear to reflect the general state of economic activity, thus both are considered to represent a separate dependent variable in a set of equations, each equation comprising common observed and common unobserved regressors. Moreover, the SUR allows for additional, equation specific determinants to be included in the system. Given that our dataset is comprised as a panel of three countries, initially we could employ a panel SUR with *3xT* equations, with two equations for precautionary and involuntary excess liquidity and one equation for the output gap. However, based on Maddala's (1997) argument discussed in the previous section, the SUR estimation method enjoys the virtue of estimating different slope coefficients for each cross-sectional unit. Estimating the model within a panel framework *3xT* would impose restrictions on the investigation of the heterogeneity among countries in our sample. With this approach we would have three equations estimating the same panel dataset. So, instead of using a panel SUR (xtsur) with *3xT* equations, we will disaggregate the estimation procedure into *9XT* equations, where we lay out 9 equations, 3 for each of the three countries for all three dependent variables (precautionary excess liquidity: 3xEL^P, involuntary excess liquidity: 3xEL^I, the output gap: 3xOG). We will not use a 3xT equations system in a panel framework providing these two arguments against:

a) A *3xT* approach imposes slope homogeneity across countries; the only variation would be the fixed effects in the dummy variables or in the error term, whereas estimating separately (country by country with 9 equations) allows complete slope heterogeneity across countries. This way, with a *9xT* structure, instead of restricting the model in a panel framework where the beta coefficient would be the same for all the countries, we enable a more flexible and technically less demanding structure. More information would be extracted, and the residuals in the SUR would pick up not only the time-invariant factors, but also the time-variant ones (Maddala, 1997).

b) Because of the relatively long T-periods, dynamic modelling in a panel framework due to serial correlation is massively complicated either with a FE or RE approach – both inconsistent when a model is dynamised. Therefore, the SUR model estimated is static.

In other words, we model the same model but within a more flexible structure, which enables maximising the heterogeneity between parameters, this way maximising the advantages of the SUR approach.

Regarding all the single equations as potentially part of a multiple-equation system, the 9 by T equations to be estimated simultaneously as a system of equations via the SUR method are presented as below:[113]

Czech Republic:

1) $EL^{pC}_t = \alpha^C + \alpha^C_1 RR_t + \alpha^C_2 i^D_t + \alpha^C_3 Eq_t + \alpha^C_4 gdpg1_t + \alpha^C_5 VolC_t + \alpha^C_6 VolD_t + \alpha^C_7 VolG_t + \alpha^C_8 npl_t + \alpha^C_9 d2005 + \alpha^C_{10} d2006 + \alpha^C_{11} d2007 + \alpha^C_{12} d2008 + \alpha^C_{13} d2009 + \alpha^C_{14} d2010 + \alpha^C_{15} d2011 + \alpha^C_{16} d2012 + \alpha^C_{17} d2013 + \alpha^C_{18} dGFC + \varepsilon^C_t$ (5.5)

2) $EL^{IC}_t = \beta^C + \beta^C_1 DEP_t + \beta^C_2 i^L_t + \beta^C_3 CREDF_t + \beta^C_4 LEG_t + \beta^C_5 FDI_t + \beta^C_6 CREDLEVEL + \beta^C_7 d2005 + \beta^C_8 d2006 + \beta^C_9 d2007 + \beta^C_{10} d2008 + \beta^C_{11} d2009 + \beta^C_{12} d2010 + \beta^C_{13} d2011 + \beta^C_{14} d2012 + \beta^C_{15} d2013 + \beta^C_{16} dGFC + e^C_t$ (5.6)

3) $OG^C_t = \delta^C + \delta^C_1 CREDF_{t-2} + \delta^C_2 OGEU_t + \delta^C_3 GOVEXP_t + \delta^C_4 d2005 + \delta^C_5 d2006 + \delta^C_6 d2007 + \delta^C_7 d2008 + \delta_8 d2009 + \delta^C_9 d2010 + \delta^C_{10} d2011 + \delta^C_{11} d2012 + \delta^{CC}_{12} d2013 + \delta^C_{13} dGFC + \nu^C_t$ (5.7)

[113] To distinguish between equations of each country, the parameters of the respective countries have been subscripted: 'C' for Czech Republic, 'E' for Estonia and 'K' for Kosovo.

Estonia:

4) $EL^{pE}_t = \alpha^E + \alpha^E_1 RR_t + \alpha^E_2 i^D_t + \alpha^E_3 Eq_t + \alpha^E_4 gdpg1_t + \alpha^E_5 VolC_t + \alpha^E_6 VolD_t + \alpha^E_7 VolG_t + \alpha^E_8 npl_t + \alpha^E_9 d2005 + \alpha^E_{10} d2006 + \alpha^E_{11} d2007 + \alpha^E_{12} d2008 + \alpha^E_{13} d2009 + \alpha^E_{14} d2010 + \alpha^E_{15} d2011 + \alpha^E_{16} d2012 + \alpha^E_{17} d2013 + {}_{18} dGFC + \varepsilon^E_t$ (5.8)

5) $EL^{IE}_t = \beta^E + \beta^E_1 DEP_t + \beta^E_2 i^L_t + \beta^E_3 CREDF_t + \beta^E_4 LEG_t + \beta^E_5 FDI_t + \beta^E_6 CREDLEVEL + \beta^E_7 d2005 + \beta^E_8 d2006 + \beta^E_9 d2007 + \beta^E_{10} d2008 + \beta^E_{11} d2009 + \beta^E_{12} d2010 + \beta^E_{13} d2011 + \beta^E_{14} d2012 + \beta^E_{15} d2013 + \beta^E_{16} dGFC + e^E_t$ (5.9)

6) $OG_t^E = \delta^E + \delta^E_1 CREDF_t + \delta^E_2 OGEU_t + \delta^E_3 GOVEXP_t + \delta^E_4 d2005 + \delta^E_5 d1998 + \delta^E_6 d1999 + \delta^E_7 d2000 + \delta^E_8 d2001 + \delta^E_9 d2002 + \delta^E_{10} d2003 + \delta^E_{11} d2004 + \delta^E_{12} d2005 + \delta^E_{13} d2006 + \delta^E_{14} d2007 + \delta^E_{15} d2008 + \delta^E_{16} d2009 + \delta^E_{17} d2010 + \delta^E_{18} d2011 + \delta^E_{19} d2012 + \delta^E_{20} d2013 + \delta^E_{21} d2014 + \delta^E_{22} dGFC + v^E_t$ (5.10)

Kosovo:

7) $EL^{pK}_t = \alpha^K + \alpha^K_1 RR_t + \alpha^K_2 i^D_t + \alpha^K_3 Eq_t + \alpha^K_4 gdpg1_t + \alpha^K_5 VolC_t + \alpha^K_6 VolD_t + \alpha^K_7 VolG_t + \alpha^K_8 npl_t + \alpha^K_9 d2005 + \alpha^K_{10} d2006 + \alpha^K_{11} d2007 + \alpha^K_{12} d2008 + \alpha^K_{13} d2009 + \alpha^K_{14} d2010 + \alpha^K_{15} d2011 + \alpha^K_{16} d2012 + \alpha^K_{17} d2013 + {}_{18} dGFC + \varepsilon^K_t$ (5.11)

8) $EL^{IK}_t = \beta^K + \beta^K_1 DEP_t + \beta^K_2 i^L_t + \beta^K_3 CREDF_t + \beta^K_4 LEG_t + \beta^K_5 FDI_t + \beta^K_6 CREDLEVEL + \beta^K_7 d2005 + \beta^K_8 d2006 + \beta^K_9 d2007 + \beta^K_{10} d2008 + \beta^K_{11} d2009 + \beta^K_{12} d2010 + \beta^K_{13} d2011 + \beta^K_{14} d2012 + \beta^K_{15} d2013 + \beta^K_{16} dGFC + e^K_t$ (5.12)

9) $OG^K_t = \delta^K + \delta^K_1 CREDF_t + \delta^K_2 OGEU_t + \delta^K_3 GOVEXP + \delta^K_4 d2005 + \delta^K_6 d2006 + \delta^K_7 d2007 + \delta^K_8 d2008 + \delta^K_9 d2009 + \delta^K_{10} d2010 + \delta^K_{11} d2011 + \delta^K_{12} d2012 + \delta^K_{13} d2013 + \delta^K_{14} dGFC + v^K_t$ (5.13)

Country subscripts are omitted to avoid clutter. Correlation coefficients (ρ) are calculated for each pair of error terms across the nine equations in the system.

In the previous chapter with the bank-level data, one of the variables also included in the precautionary excess liquidity equation was the real growth rate of the next year (t+1), to approximate the banks' expectations for future growth. In the bank-level model there were no reverse causality (endogeneity) issues involved on the grounds that future growth can influence a single bank to decide on the lending level and thus excess reserve holdings, but a single bank's lending may not impact the real growth rate of a country. In the current chapter where macro-level data of all the banks in a country are being used, due to possible reverse causality implications, this variable was excluded from the model.

In equations (5.5) to (5.13), $EL^I{}_t$ and $EL^p{}_t$ stand, respectively, for the ratio of involuntary and precautionary excess liquidity to total assets in each country in time *t*. The rationale for including two sets of excess liquidity equations is explained in section 4.3.1 , where excess liquidity is presented as a phenomenon reflecting both the supply side (risk perception) and the demand side (involuntary build-up) factors. Thus, we also follow the same procedure when estimating the macro-level equations. In short, precautionary excess liquidity (subscripted with 'p', i.e. $EL^p{}_t$) encompasses regulatory, banking and volatility (riskiness perceived) variables, whereas the involuntary excess liquidity (subscripted with 'I', i.e. $EL^I{}_t$) encompasses mainly factors that lead to involuntary build-up of excess reserves, like the public's decisions to deposit, investors' deposits or the strength of the legal rights variables. In involuntary excess liquidity equations an interaction variable between credit and a level shift dummy was added –*Credlevel*– to account for a possible behaviour of this variable post GFC. The δ is the constant term, δ_1 to δ_{24} are the parameter estimates of the independent variables, and ε_t, e_t , v_t are the error terms.

The inclusion of the output gap equation in the system of equations is novel, because studies that have investigated the determinants of excess liquidity did not consider a possible endogenous relationship between the two (for example Saxegaard, 2006). On the other side, theoretical guidance in setting up an output gap equation is both limited and vague. However, from basic theoretical knowledge, initially we know that the output gap has been commonly used as a proxy for the demand side of the economy (Gordon, 1997). Thus, as a variable reflecting the state of the aggregate demand in the economy, it should include all the factors comprising a full aggregate demand identity. Cameron (2005, p. 319) and Kennedy (2009, p. 171) argue that a system equation approach may also include identities (e.g. $Y \equiv C + I + G + NX$). Furthermore, as argued in section 5.6.2., the system of equations should be fully specified, otherwise the misspecification of one equation (e.g. omitting a relevant variable) will affect all the structural parameters (Kennedy, p. 180). Therefore, with a purpose of setting up a fully specified aggregate demand equation and following Moinescu (2012)[114] the third (5.7), sixth (5.10) and ninth (5.13) output gap equations are specified as follows.

[114] Moinescu (2012) sets out the output gap equation is: OG = α + Credit Flow + GAP(Eurozone) + LTY + FE. The credit flow is a proxy for consumption and investments, the output gap of Eurozone is a proxy for external demand and the long-term bond yields represent a fiscal measure. The author uses annual data from 2000 – 2011 and estimates the model with panel fixed effects method.

The OG_t stands for the ratio of output gap to GDP in each period of the respective country; $CREDF_{t-2}$ denotes the credit flow, which is based on the credit accelerator theory (Bernanke, 2007) and feeds consumption and investment. Suspecting endogeneity between *OG* and *CREDF*, the second lag[115] of the credit flow will be used on the grounds that lagged values of endogenous variables are treated as predetermined, because for the determination of the current period's values of endogenous variables, they are given as fixed or determined outside of the system. This is because the lagged endogenous variables are not simultaneously determined in the current time period (Studenmund, 2006, p. 476). [116]

As Kennedy (2009, p. 181) argues, the use of endogenous variables creates a reduced form estimates that are biased but asymptotically unbiased (assuming that errors are not autocorrelated). However, that author argues that "this is not a concern in the structural simultaneous equations, because all the estimates are biased anyway; they are chosen on the basis of their asymptotic properties" (p. 181); Given that some of the countries in our sample (Czech Republic and Estonia and Kosovo) are EU countries and Kosovo has very strong trade relationships with the EU (see discussion chapter 1), the output gap of the European Union divided by GDP - $OGEU_t$ - is taken to proxy the external sector in our model (foreign demand) and is considered as exogenous. Moinescu (2012) has used the OECD estimated output gap for the Eurozone; however, the OECD measure is of annual data only, whereas our sample data are of quarterly frequency. Given the lack of a publicly available measure for the quarterly output gap of Eurozone, we have estimated this unobserved component via the Hodrick-Prescott filter. Moinescu (2012) uses the long-term interest rate (bond yield) as a proxy for the fiscal stance. However, given that the issuance of the government securities in Kosovo started only from 2012, we will not be able to utilise this measure. Instead, the government (fiscal) expenditure -$GOVEXP_t$ will be used. In addition, to control for possible correlated cross-country shocks we include the year dummies *d2005- d12014* (d2004 set as a benchmark). To account for possible structural breaks in the series due to the incidence of the

[115] In econometric literature the second lag is the lag usually taken when one wants to attenuate endogeneity and make the variable predetermined. However, the exact rank of the lag will be later determined based on the empirical tests.

[116] Studenmund (2006, p. 476) argue that the lagged endogenous variables have more in common with the exogenous variables, thus are called predetermined variables. "The predetermined implies that exogenous and lagged endogenous variables are determined outside the system of specified equations or or prior to the current time period. Therefore econometricians tend to speak in terms of endogenous and predetermined variables when discussing simultaneous equations systems".

financial crisis, an intercept dummy - *GFC* - was also included.[117] The δ is the constant term, whereas the δ_1 to δ_n are the parameter estimates of the independent variables. The v_t is the error term.

Finally, the ρ term in the SUR framework measures the extent of correlation between the equations. If the errors across equations of output gap and excess liquidity are correlated, then we should be able to improve our estimators by taking account of them. The non-zero covariance of the respective error terms reflects the idea that unobserved variables are shared between these errors. Although, we do not necessarily observe these variables we can allow for their influence by correlating the error terms, which is like acknowledging the existence of some spuriousness[118] (Acock, p. 123). When a correlated error is included, then the coefficient on the estimates of *EL* and *OG* will most likely be reduced. This is because, a part of the relationship between indicators of excess liquidity and output gap is because of the influence of common (although unobserved) variables. The intuition behind a joint system equation mirrors the intuition of a single equation approach in the presence of serial correlation (If *e* is correlated with *u*, then knowledge of *u* can help reduce the predicted value of *e*).Without allowing for the errors to be correlated, we would most likely have larger coefficients but a relatively poor fit to our model (Acock, 2013).

After discussing the preferred estimation approach, a short discussion of the data used will follow.

5.7. Interpretation of diagnostics

The estimation strategy goes from a general (unrestricted) model as presented in equations 5.5 to 5.13 in section 5.6.3, to a more specific version in order to select the most parsimonious model. In order to select the most parsimonious model we have performed a number of preliminary regressions. Starting from the most general model, given the theoretical arguments previously discussed as well their possible endogeneity implications; we have obtained the following results.

[117] The GFC here represents only the general form of a structural break capturing the shift in the system after the GFC. In the next section, before estimating the results, the unit root tests with several structural breaks for each series will be presented, and based on the tests and significance we will decide on the type of the dummy variable: one time break (1=2008Q4 for CZ and EE, 2009Q4 for KS) or a n intercept dummy, zero until 2008Q4-the break, 1 after the break.

[118] Spuriousness means that a third variable explains the relationship between the two variables (Acock, 2013, p. 123)

Initially, the fully specified SUR model with all the explanatory variables, interaction variable between the credit to GDP and the level-shift dummy all, year dummies (except the first) as well as all three possible breaks (pulse dummy, level-shift dummy and trend-shift dummy) were included in all the nine equations of the model. Secondly, the insignificant break dummies as well as the insignificant interaction terms were excluded and the model was re-estimated.

In addition, as discussed in section 5.6.3, the real growth rate used in the bank-level model (previous chapter) was used with a lead to approximate the banks' growth expectation in the precautionary excess liquidity. In the bank-level model we argued that future growth may impact banks' decisions to lend, however one bank cannot impact the economy-wide growth of the next period. Nevertheless, in the current macro-level model there may be a possible a simultaneity problem, because all banks together can impact future growth. Thus, instead of using the real growth rate with a lead, we have included it with a lag, on the grounds that past growth can impact current banks' decisions on how much to lend and how much to hold as excess reserves, but current excess liquidity cannot impact past growth. This variable could no longer be used as a proxy for growth expectations, given the absence of a theoretical rationale. This variable turned out as statistically insignificant in the precautionary excess liquidity equation in the main SUR specified model, as well as in three other models estimated for robustness checks. Furthermore, an F-test for the joint significance of this variable in all three countries indicated that it was jointly insignificant at conventional levels of significance. Consequently, the results indicated that this variable does not have significant explanatory power and thus it was excluded from the model. This variable, however, proved significant in one model specification for robustness check, the model including a longer series of data with Czech and Estonia, where it was kept in the model.

The results from the final model specification (presented in table 5.8 and in appendix 5.2) indicate that the SUR model can significantly explain some of the determinants of excess liquidity and the output gap, especially for Kosovo and Estonia, though less so for Czech Republic. The results for the overall country-specific equations indicate that all equations are statistically significant at 1 percent level of significance (table 5.8 and in appendix 5.2). This indicates that the model has strong explanatory power, even though, as will be seen and explained in the next section, the statistical significance of the separate variables are weak or insignificant. However, in order to check the joint significance of each explanatory variable in all three countries, an F-test was performed. Regardless of the fact that separately many of the explanatory variables are insignificant; the joint-significance results indicate that 10 out of the 13 regressors in the model are jointly significant at conventional levels of significance (Appendix 5.2). This is a sign of multicollinearity in our data, an issue which will be further elaborated at the end of the following section.

In order to examine whether there is some efficiency gain from employing the SUR method, the Breusch-Pagan test that tests the contemporaneous covariance independence between the error terms was employed. The null hypothesis of zero contemporaneous covariance dependence between the errors from each equation can be rejected at 1 percent level of significance (table 5.8 and in appendix 5.2). Thus there is evidence in support of contemporaneous cross-sectional correlation among the error terms and, hence, there is an efficiency gain from employing the SUR method. Additionally, the correlation matrix of the Breusch-Pagan test provides evidence in support of the idea that there are common and linked systematic unobservables associated with both, excess liquidity and the output gap.

Table 5.8. Estimation results of the final SUR model specification, 2004Q1 – 2013Q4

Dependent: Precautionary EL/Assets				Dependent: Involuntary EL/Assets				Dependent: Output Gap/GDP			
Country/Equat.	CZ (1)	EE (4)	KS (7)		CZ (2)	EE (5)	KS (8)		CZ (3)	EE (6)	KS (9)
Variables	del_1	del_2	del_3	Variables	2del_1	5del_2	8del_3	Variables	og_1	og_2	og_3
L.rrr	-0.289	-0.516*	-0.012	*L.Dep*	-0.094**	-0.073	-0.012	*L.Cred*	0.005	-0.005	-0.299**
	(1.061)	(0.278)	-0.079		-0.048	-0.049	-0.045		-0.007	-0.012	(0.128)
L.idep	-0.311	-0.790	-1.66***	*L.Govdep*	-0.014	-0.035	0.108	*L.Ogeu*	0.325***	0.755***	-0.393
	(0.933)	(1.142)	(0.625)		-0.017	-0.038	-0.073		-0.066	(0.122)	(0.291)
L.Volcred	-0.068	1.843	0.366	*L.Cred*	-0.016	0.047	0.195	*L.Govex*	0.0234	0.007	-0.009
	(0.475)	(6.146)	(0.291)		-0.01	-0.032	(0.135)		-0.044	-0.015	-0.022
L.Voldep	-0.054	0.141	-1.444**	*L.iloan*	-0.300	-0.207	-0.336	*d2005*	1.031***	0.553	-0.303
	(1.430)	(0.753)	(0.651)		(0.721)	(0.201)	(0.308)		(0.201)	(0.434)	(1.027)
L.Volgdpg	0.003	-9.051	0.904**	*L.Fdi*	0.349***	-0.189***	0.0101	*d2006*	2.325***	1.236***	0.413
	-0.041	(20.68)	(0.362)		(0.125)	-0.036	(0.0158)		(0.208)	(0.443)	(1.032)
L.Npl	0.072	-2.89**	0.386*	*L.Rle*	14.51**	14.23	55.96**	*d2007*	2.778***	0.445	3.194***
	(0.235)	(1.351)	(0.229)		(6.335)	(16.08)	(24.70)		(0.254)	(0.527)	(1.212)
L.Eq	-0.036	-0.176	-0.82***	*L.CredLevel*	-	1.980	-0.778***	*d2008*	2.682***	-2.776***	5.847***
	(0.180)	(0.295)	(0.279)		-	(2.049)	(0.240)		(0.276)	(0.572)	(1.313)
d2005	0.198	0.188	-0.779	*d2005*	-0.309	2.347*	2.848**	*d2009*	1.920***	-6.978***	3.720***
	(0.491)	(1.659)	(0.981)		(0.448)	(1.369)	(1.413)		(0.204)	(0.444)	(1.038)
d2006	-0.405	-2.401	-0.341	*d2006*	-0.937**	-2.302**	1.398	*d2010*	2.343***	-5.998***	3.650**
	(0.603)	(1.883)	(0.944)		(0.433)	(1.125)	(1.117)		(0.306)	(0.626)	(1.481)
d2007	-0.526	0.546	-1.796*	*d2007*	-1.418***	-0.526	-0.927	*d2011*	1.632***	-5.634***	5.559***
	(0.626)	(2.411)	(0.934)		(0.458)	(1.246)	(0.976)		(0.251)	(0.525)	(1.226)
d2008	1.864***	4.483*	-0.576	*d2008*	1.241***	0.412	-1.073	*d2012*	0.461*	-5.332***	5.060***
	(0.688)	(2.663)	(0.942)		(0.441)	(1.182)	(1.078)		(0.269)	(0.555)	(1.345)
d2009	0.895	4.170	1.534	*d2009*	0.620	-0.688	1.453	*d2013*	-0.429	-3.993***	1.641
	(0.556)	(2.711)	(0.970)		(0.437)	(0.969)	(1.363)		(0.304)	(0.634)	(1.518)
d2010	0.267	2.552	-0.002	*d2010*	0.259	2.036**	2.268	Constant	-1.278	3.638***	-4.861***
	(0.559)	(1.926)	(0.971)		(0.413)	(1.010)	(1.667)		(1.859)	(0.354)	(0.826)
d2011	0.876	-1.075	-1.011	*d2011*	0.950**	4.350***	1.035				
	(0.538)	(3.081)	(0.975)		(0.462)	(0.913)	(1.751)				
d2012	0.514	-7.750**	0.585	*d2012*	0.493	0	0.978				
	(0.541)	(3.872)	(0.948)		(0.422)	(0)	(1.565)				
d2013	1.113*	-8.330**	1.478	*d2013*	1.087***	-1.683*	1.281				
	(0.625)	(3.682)	(0.968)		(0.411)	(0.950)	(1.519)				
Constant	-	6.767*	-	*Dpulse*	0.800**	-0.055*	-				
	-	(3.755)	-		(0.406)	-0.03	-				
				Dlevel	-	5.682***	3.681**				
					-	(1.952)	(1.576)				
				Constant	-1.122**	2.729***	-2.762**				
					(0.438)	(1.007)	(1.170)				
Observations	38	38	38		38	38	38		38	38	38
R-squared	0.642	0.492	0.551		0.721	0.740	0.489		0.980	0.987	0.821
RMSE	0.534	1.867	1.105		0.471	1.335	1.179		0.234	0.508	1.157
F-stat for the joint significance of the country specific equations											
	78.4***	36.6***	58.9***		110.65***	128.93***	42.81***		1891.33***	2911.95***	180.2***
Breusch-Pagan test for the contemporaneous covariance independence between the error terms chi2 = 116.451; p-value = 0.000											

Note: ***Significant at 1% l.s., ** significant at 5% l.s., *significant at 10% l.s.

Source: Author's own calculations performed in Stata 12.

5.7.1. Interpretation of the results

Box 5.1. The interpretation of parameters in differences

As discussed earlier, one of the requirements of the SUR modelling technique is stationarity in the data. The integrated data of order one, as indicated by several unit root tests, were then first-differenced. The theory discussing the determinants of excess liquidity is in terms of the levels of variables. In order to examine whether the interpretation of the estimated parameters in first-differences differs from those in levels, the following example is presented below:

A: $Y_t = \hat{c} + \beta X_t + \varepsilon_t$
B: $Y_{t-1} = \hat{c} + \beta X_{t-1} + \varepsilon_{t-1}$
C: A – B: $Y_t - Y_{t-1} = (\hat{c} - \hat{c}) + \beta(X_t - X_{t-1}) + \varepsilon_t - \varepsilon_{t-1}$ or
C.1. $\Delta Y_t = \Delta \beta X_t + \varepsilon_t$

Equation A represents an equation where the beta parameter - β - is estimated from the variable X in levels, and equation B is the lagged version of equation A. Equation C results from the first-differencing of the equation A, which is equation A minus equation B. As can be noted, the beta parameter is the same in the levels as in the differenced equation. This means that if the relationship in the levels is positive (negative), that relationship should also carry over to the differences. This is because, if two variables are positively correlated over time, their mean growth rate should also be positively correlated, even if some periods are characterised by negative values. Nevertheless, when data are differenced there is some loss of information, therefore it is harder to estimate the beta coefficient precisely. Thus, even though the data are differenced, the estimated coefficient sign is expected to be the same as in the levels, because the same parameter is being estimated, i.e. the β parameter represents a constant elasticity. The same principle is also followed in other estimation techniques, such is the difference Generalised Method of Moments (GMM) estimator, in which the group specific effects are differenced away to overcome endogeneity or correlation between some variables and the group specific effects. Therefore, interpretation of the estimated coefficients on the differenced variables below will be discussed as in levels.

As presented in table 5.8, there are differences in the estimated signs as well as the sizes of the coefficients across countries. This indicates that the three countries in our sample with different levels of economic and financial development tend to exhibit heterogeneous behaviour regarding excess liquidity and the output gap.

With regard to the determinants of precautionary excess liquidity, more variables appear to be significant in Estonia and Kosovo, compared to the Czech Republic (where none of the explanatory variables turned out to have a statistically significant impact on the accumulation of the precautionary excess liquidity). As expected, the reserve requirement ratio (*RRR*) appears with a negative coefficient for all three countries, which is in line with the theory suggesting that rising the reserve ratio (which are typically non-remunerated in these countries) also raises the overall cost of holding reserves and thus may induce banks to reduce their precautionary holdings of excess reserves. However, this variable is significant only in the case of Estonia. This may be explained due to the fact that this ratio was unchanged throughout the period for the Czech Republic and Kosovo, meanwhile there were two policy induced changes in the reserve requirement ratio in Estonia; the first one associated with an increasing concern with maintaining the stability of the financial system is likely to have caused increases in the required reserve ratio for prudential reasons at the beginning of 2007 and the second one more of a gradual shift towards converging with the ECB reserve ratio of 2 percent.

The deposit interest rate (*idep*), approximating the funding costs of the banks, also appears to uniformly reduce the precautionary excess liquidity, however it is statistically significant for Kosovo only and with a relatively higher magnitude. A one pp increase in the deposit interest rate would reduce the change in precautionary excess liquidity by 1.6 pp. This is not in line with orthodox theory suggesting that, other things being equal, when funding costs raise banks are expected to hold larger amount of excess reserves to prevent liquidity shortages. As a matter of fact, banks in European transition economies like our sample countries, especially Kosovo, rely on the domestic deposits as their main funding source, especially having in mind the negligible role of the inter-bank market. While deposit interest rates are relatively lower in the Czech Republic and Estonia, in Kosovo they sometimes reached 5 percent and thus may be considered to be relatively costly funds. Thus, orthodox theory may not necessarily explain well the behaviour of excess liquidity with regard to the funding costs in European transition economies, because our result indicate the opposite: if the cost of acquiring liabilities increases, then the quantity acquired should fall because banks would rather use up their own excess funds to finance loans or other investments.[119] The insignificant and rather low

[119] Conversely, in developed European countries, e.g. UK, especially during the post-crisis period, due to quantitative easing central banking policies, commercial banks appeared to have considerable amounts of excess liquidity, because it was offered almost for free, thus the opportunity cost of holing it in this case is almost zero.

coefficients on changes of deposit interest rates for Czech Republic and Estonia may reflect the fact that these rates are quite low in these two countries, as amongst other factors they also reflect the convergence process of interest rates in the Eurozone.[120] This may imply that deposits interest rates reflect rather low funding costs. The volatility of credit (*Volcred*) appears with an insignificant impact on the precautionary excess liquidity for these three countries, because it may not be a good proxy for perceived banks' risk for our sample countries. The volatility of deposits (*VolDep*) is highly significant only for Kosovo. The negative coefficient sign is not in line with theory suggesting that as volatility of deposits increases banks act to insure themselves against shortfalls in liquidity by increasing the precautionary excess liquidity. Even though this finding is not in line with prior expectations, Saxegaard (2006) also reports a negative coefficient on this variable. One pp change in the volatility of deposits, ceteris paribus, would reduce the change in precautionary excess liquidity by 1.4 pp, implying that banks in Kosovo will have to use their own funds when deposits are more volatile and less predictable. The volatility of the growth (*Volgdp*) rate is only significant in the case of Kosovo and the coefficient has the expected sign. A one unit increase in the volatility of growth rate, ceteris paribus, would increase the insecurity regarding the economic performance (which is associated with the loan repayment expectations), thus increase the change in precautionary excess liquidity by 0.9 pp.

The changes in non-performing loans ratio (*NPLs*), approximating a risk measure for the banking business, is statistically significant in the Estonia and Kosovo cases, nevertheless, the respective coefficients are with opposing signs. As expected, a one pp increase in the change of NPL rate would, other things being equal, induce banks to increase the change in precautionary excess liquidity ratio in Kosovo by 0.4 pp. In Estonia, the change in the NPL ratio would, other things being equal, reduce the change in precautionary excess liquidity by 2.9 pp. This finding is not in line with expectations, since higher NPLs reduce the expected incomes of the banks thus induce them to lend less and hold more excess reserves. One explanation may be weak managerial skills in liquidity management as well as a failure to properly monitor and screen the loan applications. Secondly, regardless of the rising NPLs which may also reflect past decisions on the issuance of the loans that appear in current

[120] The Eurozone countries like Estonia and those within EU that are part of the European Monetary Union are in a converging process of monetary policy instruments. From 2008 (2% deposit facility rate, 2.5% refinancing operations rate) onwards the ECB's key interest rates have been constantly dropping, reaching the historically lowest rate in 2014 (-0.2% for deposit facility and 0.05% for refinancing operations) with the purpose of offering quantitative easing in member countries that have been undergoing a rather long and hard recuperating economic process since the onset of the global financial crisis.

financial reports of the banks (e.g. pre-crisis issued loans), the deposit flows have been lower post-crisis, so banks have had to use up their own excess funds for daily transactions. The changes in the ratio of equity to assets (*Eq*) appear to uniformly induce banks to hold less precautionary excess liquidity. This finding is in line with expectations since equity already serves as a precautionary buffer, albeit for longer term obligations. Meanwhile, additional equity requirements are policy-induced restrictions on the banks, i.e. restricted funds that banks cannot use for doing business, so they also represent additional costs. Therefore, when required to increase the equity to assets ratio, banks may want to hold less precautionary excess reserves. Finally, the year dummies controlling for the common shocks across countries are significant in the two countries hit severely by the financial crisis in 2008, namely Czech and especially Estonia, capturing the financial crisis impact in the banks of these two countries. In the year 2008, the positive year dummy coefficients indicate that banks in Czech and Estonia accumulated more precautionary excess liquidity as compared to the benchmark year of 2004. Also, years 2012 and 2013 also appear significant having in mind that they represent periods of stagnation, after a recuperating period in 2009 – 2011.

The determinants of involuntary excess liquidity seem to have a higher impact on the behaviour of excess liquidity for the Czech Republic than in the other two countries. The change of deposits to GDP ratio (*Dep*) appears to negatively impact the excess liquidity to assets ratio in all three countries, even though this impact is statistically significant only in the case of the Czech Republic. A one pp change in deposits to GDP ratio reduces the change in involuntary excess liquidity ratio by 0.09 pp. This finding is not in line with the theory suggesting that an increase in deposits is thought to reflect in an increase of involuntary excess liquidity. The government deposits to GDP ratios (*Govdep*) are statistically insignificant in all three countries. This may be due to the fact that compared to overall deposits, the government deposits are lower thus have an insignificant impact. The effect of a change in the credit to GDP ratio (*Cred*) is also statistically insignificant on the accumulation for the changes of involuntary excess liquidity to assets ratio in all three countries. Even though *Cred* is the only common variable linking excess liquidity with the output gap, its impact does not appear to be dominant in explaining the accumulation or reduction of the involuntary excess liquidity. As will be later discussed, this may partially be due to other external factors impacting the whole system and partially due to the multicollinearity problem that may be present in the data. The estimated results also indicate an insignificant impact with regard to the loan interest rate (*iloan*) all three countries.[121] The FDI to GDP ratio (*FDI*) is highly significant in Czech Republic and Estonia case, but insignificant for Kosovo. This result is to be expected, because the FDI inflows represent a more important funding source in Czech and Estonia relative to Kosovo. A one pp increase in the FDI inflows to GDP ratio would increase the changes in involuntary excess liquidity by 0.35 pp, while decrease it in Estonia for 0.19 pp, ceteris paribus. In the case of Estonia this result may be picking up mostly the after-crisis behaviour of the FDI inflows, which is the period when the FDI inflows reduced drastically and did not rise back to the pre-crisis levels.[122]

It should be noted that there is a lack of theory for the inclusion of the Rule of Law Estimate (RLE) index (or other similar proxies) in the precautionary excess liquidity and the output gap equations. Regardless though, considering the likely importance of this proxy in the transition context and the lack of more common observables in the model, we have also estimated the preferred SUR specification including the RLE index in all the equations (see Appendix 5.8). Nevertheless, as can be seen in table 5.15 (Appendix 5.8), consistent with the theory, the RLE index turns insignificant in the precautionary excess liquidity and the output gap equations. Furthermore, most of the model remains similar to the main specified SUR model, regarding the coefficient sign, size and significance. Thus, the RLE index was kept only in the main model specification. The generated results regarding the RLE indicate quite a similar impact on Czech and Estonian involuntary excess liquidity, quantitatively and qualitatively. On the other side, the impact of this index in Kosovo (Kosovo ranks relatively lower than the other two countries in this index) seems to have a quite profound impact in the Kosovo banking system. The relatively large coefficient magnitude in all three countries can be explained by the relatively small range of the RLE index (from -2.5 to +2.5). In the case of the Czech Republic, Estonia and Kosovo the RLE range is even smaller, respectively ranging from 0.74 – 1.02, 0.73 – 1.17 and -1.03 – 0.55, indicating an even smaller range than 1 in absolute

[121] It should be noted that, similarly to the analysis presented in chapter 4, we have also considered the inclusion of a squared lending interest rate variable in the model to account for a possible non-linear relationship between excess liquidity and lending rates. However, loan interest rate squared turned out to be insignificant. Saxegaard (2006) argued that the loan rate may be sticky because of imperfect information about potential new borrowers. Sticky loan rate may also be case in European transition economies, because as we noted in chapter 1, lending interest rates kept broadly to the same level for over a decade. Finally, the reason why lending interest rates squared turned out as significant in chapter 4 while insignificant in chapter 5 could also be the use of different datasets; bank-level data in the former and macro-level data in the later.

[122] To pick up the post-crisis behaviour of the FDI to GDP in involuntary excess liquidity, this variable was also interacted with the intercept shift dummy. However, this specification turned out to be statistically insignificant as a separate variable, as well as when tested for joint significance in the three countries, thus it was not pursued further.

terms. This is why an increase of 1 unit pp of RLE results is quite a large coefficient magnitude. Thus, the interpreted results for the RLE indicator will only be in qualitative terms. As proclaimed in the literature, a better perception of the extent that private agents have confidence in and abide with the legal, institutional and political framework is expected to contribute to reducing the asymmetric information problem between banks and borrowers and as such induce the banks to issue more loans and hold less excess reserves. Nevertheless, the situation in transition economies may be more complex. This is because with better institutional and legal frameworks, higher transparency may be required from borrowers during the application process, such as the proper reporting (financial statements, especially for enterprises) and tax paying evidence. Hellström (2009) finds that the financial accounting quality in transition economies is lower than in the developed countries (with special reference to the Czech Republic), which directly impacts on the investors' decisions and distribution of wealth between individuals. Based on the practitioner's knowledge, enterprises in Kosovo, which are the main borrowers from banks (around 70 percent), tend to hold two types of accounting reports: one with deflated numbers for the tax authorities, and another one with inflated profits for the banks when applying for a loan. Therefore, better institutional and judicial institutions in transition economies would require the disclosure of the information of the related parties, and given the unrealistic situation prevailing (much lower profits and thus lower repayment capabilities), would actually induce banks to increase the involuntary excess liquidity holdings even more. Thus, the positive sign of the RLE index may be sensible in our sample.

The results presented in table 5.8 indicate that the determinants of the output gap to GDP ratio act in a similar fashion in Czech and Estonia but differently for Kosovo. The changes in credit to GDP ratio appear to negatively impact the output gap in Estonia and Kosovo, however the impact is significant only in the latter. This finding is in line with the arguments of Biggs and Mayer (2013) who suggest that the business cycle and the change in credit i.e. 'credit impulse' are positively correlated. As expected, a higher credit impulse is expected to raise consumption and investment, therefore reflected in positive growth and a lower output gap. The credit impulse variable is significant only in the Kosovo case. On the other side, as expected, the results suggest that the output gap of Eurozone is highly and positively correlated with the output gap to GDP of Czech Republic and Estonia. A one pp increase in the ratio of the (below potential) output gap of the Eurozone, on average, would increase the (below potential) output gap in Czech and Estonia by 0.3 and 0.7 pp, respectively, ceteris paribus. This result is expected, given that both the Czech Republic and Estonia are part of the same system, that of the European Union (EU). Furthermore, from 2011 Estonia was also part of the Euro area. Therefore, these two countries, besides having close external sector relations with Eurozone countries, are also closely linked due to similar or converging towards the same policy frameworks regarding monetary, financial and fiscal sectors. This finding is also in line with the findings of Moinescu (2012) who also finds that the Eurozone output gap has the largest impact (and positively correlated) on these countries' output gaps.

Overall, the SUR model does not provide a fundamental variable that would affect both, excess liquidity and the output gap significantly and in the same direction. The only variable in common, the credit to GDP variable (*Cred*), is insignificant in most cases and with opposing signs in the excess liquidity and output gap equations. Given the high explanatory power of these equations but the relatively low significance of the individual variables in the model, the model seems to be more correlated via the unobserved components rather than via the observables in the model. The year dummies picking up otherwise unobserved factors impacting both excess liquidity and the output gap suggest the following pattern: for the Czech Republic, the unobserved factors prior to 2008 appear to impact the excess liquidity and the output gap in the opposite direction, whereas from 2008 onwards (except for 2013) these unobserved factors act as a system, consistently pushing the two variables to increase. In Estonia the pattern is less clear, however in 2008 the unobserved factors act in opposite directions, whereas in 2012 and 2013 these factors seem to reduce both, excess liquidity and the output gap. In the case of Kosovo, the underlying unobserved factors in most cases seem to reduce excess liquidity while inducing an increase in the output gap.

The correlation matrix of the respective equations' residuals, taken from the Breusch-Pagan test (Appendix 5.2.) is presented below:

Table 5.9. Correlation matrix of the residuals of equations - r

Countries		Czech Republic			Estonia			Kosovo		
	Equations	EL^P	EL^I	OG	EL^P	EL^I	OG	EL^P	EL^I	OG
Czech Republic	EL^P	1.000								
	EL^I	0.723***	1.000							
	OG	0.060	-0.096	1.000						
Estonia	EL^P	-0.172	-0.226	0.039	1.000					
	EL^I	-0.116	-0.292*	0.073	0.636***	1.000				
	OG	-0.112	-0.248	0.213	0.089	-0.123	1.000			
Kosovo	EL^P	-0.338**	-0.172	-0.447***	0.055	0.139	-0.315*	1.000		
	EL^I	-0.303*	-0.252	-0.148	0.090	0.088	-0.356**	0.705***	1.000	
	OG	0.032	0.160	-0.025	-0.066	-0.031	-0.52***	0.249	0.394**	1.000

Note: ***significant at 1% l.s., **significant at 5% l.s., *significant at 10% l.s.

Source: Author's own calculations performed in Stata 12.

The Breusch-Pagan test results presented in table 5.9 (appendix 5.2) provide some evidence that the residuals of both types of excess liquidity and the output gap are correlated via the unobserved factors.[123] The correlation coefficients (r) presented in table 5.8 represent the degree of linear association between two variables (equations' residuals in our case), i.e. how closely the data fit a linear relationship. Nevertheless, even though a relatively high correlation coefficient may present indicating the strength of the relationship between the two factors, like EL^P and EL^I or EL^I and OG, these coefficients should also be assessed regarding their statistical significance. This is because, the significance of a correlation coefficient ensures that the relation between a pairwise factors is not by pure chance. Thus, the large or low magnitude of the correlation coefficient is not necessarily a good indicator of the respective relation. For example, in relatively large samples (e.g. N > 100) even smaller correlation of around 0.2 may be significant (e.g. involuntary excess liquidity between the Czech Republic and Estonia), whereas in some cases with small samples, even correlations coefficients higher than 0.3 may turn out as statistically insignificant (Taylor, 1992). Following Taylor (1992), the correlation coefficient significance in our case was determined based on the sample size to calculate the degrees of freedom (in our case the degrees of freedom = 38 – 2 = 36;) and the Pearson's two-tailed critical values ($CV_{5\%} = 0.32$). As can be noticed in table 5.10, most of the correlation coefficients higher than 0.3 are usually statistically significant at conventional levels, indicating a strong and significant relationship

123 The correlation coefficient r, in absolute terms, approximately < 0.35 is generally considered to represent low or weak correlation, 0.36 < r < 0.67 as a moderate correlation, whereas > 0.90 suggest very high correlation (Taylor, 1992).

between our three factors of interest.[124] The correlation coefficients between precautionary and involuntary excess liquidity residuals appear to be consistently highly and positively correlated in all three countries. Additionally, the magnitude of the three coefficients is quite similar. The relatively high and significant correlation coefficients between the residuals of precautionary and involuntary excess liquidity in all three countries indicate that there are same unobserved factors inducing both types of excess liquidity to move in the same direction. The same direction of correlation goes with precautionary excess liquidity and the output gap in all three countries, however the coefficients are statistically insignificant, except for Kosovo, where the $r = 0.25$ is at borderline of 10 percent level of significance. This means that the same unobserved factors appear to be driving precautionary excess liquidity and the output gap in the same direction. With regard to the correlation between involuntary excess liquidity and the output gap the correlation coefficient in negative for the Czech Republic and Estonia, meaning the same unobserved factors are pushing the two factors in opposite directions. This systematic conjecture is even clearer in the Kosovo case, where the correlation coefficients are even more substantial than in the Czech Republic and Estonia. Not only the precautionary and the involuntary excess liquidity are positively and significantly correlated by the unobserved factors, but both types of excess liquidity are also positively correlated with the output gap. Thus in the Kosovo case, the same unobserved factors appear to be driving both types of excess liquidity and the output gap in the same direction. These findings are in line with the previously stated conjecture that excess liquidity does not necessarily cause the output gap. Once more the evidence points to a correlation relationship between excess liquidity and the output gap.

Despite the fact the correlation coefficient can measure the strength of a pairwise relationship between variables or in our case between two equation residuals, and that the significance of that correlation can be tested to ensure that the correlation is not accidental, these coefficients are abstract and hard to interpret. This is because, the correlation coefficient only indicates the strength of the relation between two variables and the coefficients are unit free. Although the correlation coefficients are widely used in empirical studies, Taylor (1992) suggests the use of the *coefficient of determination* is a more meaningful measure and provides more direct and precise interpretation. "The coefficient of determination (r^2) is defined as the percent of the variation in the values of the dependent variable (y) that can be 'explained' by variations in

[124] The significance of the correlation coefficient indicates how unlikely a given correlation coefficient *r* may occur, given no relationship in the sample.

the value of the independent variable (*x*)" (Taylor, 1992, p. 37). Nevertheless, it should be noted that no matter how high a correlation coefficient may be, that does not preclude a cause-and-effect relationship (Taylor, 1992).

The calculated coefficients of determinations are presented in table 5.10 below.

Table 5.10. The coefficients of determination - r^2

Countries		Czech Republic			Estonia			Kosovo		
	Equations	EL^P	EL^I	OG	EL^P	EL^I	OG	EL^P	EL^I	OG
Czech Republic	EL^P	1.000								
	EL^I	0.523	1.000							
	OG	0.004	0.009	1.000						
Estonia	EL^P	0.030	0.051	0.002	1.000					
	EL^I	0.013	0.085	0.005	0.404	1.000				
	OG	0.013	0.062	0.045	0.008	0.015	1.000			
Kosovo	EL^P	0.114	0.030	0.200	0.003	0.019	0.099	1.000		
	EL^I	0.092	0.064	0.022	0.008	0.008	0.127	0.497	1.000	
	OG	0.001	0.026	0.001	0.004	0.001	0.270	0.062	0.155	1.000

Source: Author's own calculations.

In the case of the Czech Republic, even though the correlation coefficient (r) between precautionary and involuntary excess liquidity is 0.72, the coefficient of determination $r^2 = 0.52$ indicates that only 52 percent of the total variation in the residuals of one type of excess liquidity may be explained by the variation in the other type of excess liquidity residuals. In the case of Estonia, the $r = 0.64$, whereas the $r^2 = 0.40$ indicating 40 percent explanation of variation. This pairwise correlation in the case of Kosovo is $r = 0.70$, whereas the $r^2 = 0.50$ indicating 50 percent explanation of variation; furthermore, even though the significant correlation coefficient between involuntary excess liquidity and the output gap is $r = 0.39$, the $r^2 = 0.15$ indicates that only 15 percent of the total variation in the residuals of one variable is accounted for by the variation in the residuals of the other variable.

In addition, the results from the Breusch-Pagan test can also be supported via the period dummies' effects in table 5.8 and even clearer in table E5.14 in Appendix 5.2; there are systematic unobserved period effects on both types of excess liquidity and the output gap. The period dummies serve as evidence, indicating that both types of excess liquidity and the output gap have common determinants –year dummies – picking up common global events, such as the global financial crisis. This is to be expected, given that all three countries in our sample are small open economies, and they are also prone to common external developments, which once more support the idea that they are linked via the common observed and unobserved factors. Therefore, we are dealing with a system where the linkage comes from either unobserved non-systematic events in the error term, or unobserved systematic influences captured by the period effects (black box). Nonetheless, this finding does not support any policy intervention, since the external events are not subject to domestic policy interventions. What we do not have is common observables with known and predictable effects to be reliable for policy intervention.

5.8. Robustness checks

Robustness checks of the model have been undertaken in the following ways. First, as mentioned in section 5.6.2, after a sufficient number of iterations in estimating the coefficients and their variances for each cross-sectional unit converge, then the FGLS estimator equals the Maximum Likelihood Estimator (MLE). Hence, this is one way to compare if the estimates already reported in table 5.8 estimated by the FGLS estimator are in line with the ones calculated by MLE. However, as mentioned in section 5.6.2, whether the MLE provides additional efficiency gains over the FGLS and which estimator has better asymptotic properties remain open issues. Nevertheless, this robustness check was not possible to implement since the SUR FGLS model with iterations did not converge. Therefore, instead of the SUR MLE robustness check, as mentioned in the data section, the main SUR model was estimated via the OLS equation-by-equation in order to compare the size of the coefficients between the two methods that according to the Moon and Perron (2006) are expected to be quite similar. Secondly, in order to add confidence to the empirical results given the presence of the interpolated data in the model, we have estimated five more SUR FGLS models as a robustness check: the main SUR model with HP-filtered Eurozone output gap (instead of the interpolated OECD estimate of the Eurozone output gap); the main SUR model with a minimum number of interpolated data (except for the OECD Eurozone estimate of the output gap and the Rule of Law Estimate) but with a shorter time span – 2008Q1 – 2012Q4 and a model with the Czech Republic and Estonia only with a longer series – 2002Q1 to 2012Q4. In addition, given that the unit root tests for a short time-span of series have weak power to reject the null of a unit root, even though in a loner-term the series may be stationary, thus the following robustness check of the main SUR specification model will treat the excess liquidity ratio as a stationary variable. This is because, in a longer-term period, excess liquidity expressed as a ratio may represent a random walk or a random walk with a drift, because it cannot infinitely rise or decrease. The final robustness check was performed to including the RLE index in all the model equations (Czech Republic, Kosovo, Estonia, 2004Q1 – 2013Q4).

According to the first type of robustness check, i.e. OLS equation-by-equation (see table E5.10 in appendix 5.3) the results indicate that the standard errors estimated by OLS equation-by-equation (see appendix 5.3) are higher as expected, than the ones estimated by the SUR model (see table 5.8 and appendix 5.2). This directly affects the significance of the variables as well as the overall significance of the regressions for each cross-sectional unit. In the estimates obtained by the SUR model, all of the nine overall regressions were statistically insignificant for all three countries, whereas when employing OLS equation-by-equation two out of nine for all three countries were statistically insignificant (see F-tests in table 5.9, appendix 5.3). This may suggest that there is indeed some efficiency gain by employing the SUR model. Regarding the estimated size of the coefficients for each equation individually they are qualitatively similar to the ones obtained by SUR model with the FGLS estimator (see table 5.8 in section 5.7 and table E5.10 in appendix 5.3), but quantitatively larger. The larger coefficients generated by the OLS compared to the ones in SUR are to be expected. This is because, conversely to the OLS approach, the system SUR modelling approach estimates the coefficients by also accounting for the unobserved influences in the model, thus generating lower coefficients.

In the second type of robustness check the OECD estimate of the Eurozone output gap (with interpolated data for the first three quarters) in the main SUR model specification was substituted with the HP-filtered estimate of the Eurozone output gap (author's own calculations with no interpolated data). The correlation between the HP-filter estimate and the OECD estimate of the Eurozone output gap indicate a relatively large correlation between the two of 0.63 (Table E5.11 and Appendix 5.4). The overall results are quite comparable qualitatively and quantitatively, with small differences. The estimated coefficients of the HP-filtered Eurozone output gap has a positive and highly significant impact in Czech and Estonia, and a negative and significant impact in Kosovo. These results are quite similar to the main SUR model. However, with the HP estimate, the Eurozone output gap has a significant impact on Kosovo, too. Even though the two models with different estimates of the Eurozone output gap generated quite similar results, the main SUR specification was chosen with the OECD estimate, because it generated a slightly better model fit.

Trying to estimate the SUR model with the minimum number of interpolated values, the third type of robustness check takes place with all three countries but with a shorter series span. This model covers the period around the onset of the global financial crisis 2008Q1 to 2012Q4, so no break dummies are included. This model, in addition to the robustness check of the main model may also serve as a model depicting any different behaviour in all the structural variables in our model, as compared to the period before 2008. In this SUR model, even though the results in most of the cases are relatively similar, some of the coefficients are slightly higher in some cases (Table E5.12 in appendix 5.5). Additionally, a few variables like *RRR (switched sign), VolCred* and *Govdep (switched sign)* for the Czech Republic, while for Kosovo or *idep, VolGDP Cred* and *RLE (switched sign)* are now statistically significant. Another example is the Eurozone output gap, which is statistically significant in all other model specifications (main model and robustness checks), nevertheless in this specification it is insignificant for the Czech Republic.

When comparing the results of the model with a longer period covering 2002Q1 – 2012Q4 it appears that there are more differences in the results (see table E5.13 in appendix 5.6). Considering the overall statistical significance of the regressions for each individual equation, again all six equations in the two countries are statistically significant. This model is different from other model specifications since it does not include Kosovo and includes the lagged growth rate as an explanatory variable. The latter is because the growth rate variable was statistically significant only in this model specification; in the F-test for joint statistical significance in both countries and separately as an explanatory variable in the case of Estonia. A one pp increase in the previous period real growth rate would lead to 0.3 percentage change increase in the precautionary excess liquidity, which is contrary to the prior expectations. Further, when covering a longer period, the non-performing loans ratio was statistically significant in this model only for Czech Republic and has the expected sign. A one pp increase in the non-performing loans ratio, would other things being equal, lead to an increase in the change precautionary excess liquidity for 0.3 pp. Furthermore, the credit to GDP ratio is also significant for the Czech Republic in this model specification. A one pp increase in the change of the credit to GDP in the previous quarter would reduce the change in involuntary excess liquidity by 0.03 pp.

Having in mind the discussion in section 5.5 regarding the generated results from the unit root tests, another set of robustness check treating excess liquidity ratio as stationary has been implemented. This is because, our sample data, even though in quarterly frequency covers a relatively short time-span. The short time-span of the data weakens the ability of unit root tests to reject the null of a unit root. In addition, the fact that excess liquidity is expressed as a ratio (which is bounded between a certain range) which would lead us to believe that it may not represent a random walk (with a drift) process, as after a while the series should be mean reverting, in comparison to the main model specification, the results presented in table E5.14 indicate some differences in the magnitude and significance but also a switch in the sign of some coefficients. Significant variables that have switched their coefficient sign include the *RRR*, *VolDep*, *Dep*, and *FDI* for the Czech Republic; *RRR*, *Dep*, *Govdep*, and *RLE* for Estonia, and for Kosovo, *RRR* and *VolGDP*. Most importantly, the year dummies picking up the unobserved factors impacting the output gap and excess liquidity display a consistent pattern. In this model specification, the same underlying but unobserved factors inducing the reduction of both types of excess liquidity, most of the time lead to an increase in the output gap. The opposite holds for the case of Estonia. In the Kosovo case, the same underlying factors reducing the accumulation of both types of excess liquidity lead to a reduction of the output gap, too.

After considering the main specification results and all the sets of other robustness check models a few final observations should be made. Even though the output gap equations for all three countries are quite robust, the estimated results for the precautionary and involuntary excess liquidity change in some cases. On the other side, a few variables tend to change significance and switch coefficient signs in different model specifications. The overall model F-statistic is highly significant suggesting a high explanatory power, while individual explanatory variables tend to have low t-statistics, though a consistent coefficient sign is found for some variables (e.g. *Cred*, *Dep* or *RLE*). These are all indicators that mulitcollinearity may be an issue in our data (see variance inflation factor checks in appendix 5.3), which may be causing the pervasive coefficient signs in some of the variables. However, the multicollinearity issue cannot be addressed in small sample sizes like ours.

5.9. Conclusion

Compared to the large set of previous studies investigating the output gap and the few ones investigating the determinants of excess liquidity, this chapter is novel in two respects. Firstly, a new conceptual relationship between the output gap and excess liquidity is introduced and secondly, the application of a method not previously used to account for potential endogeneity between the two, by allowing both excess liquidity and the output gap to be determined by a range of similar and different underlying determinants.

Initially, the main research question addressed in this chapter was ‘does the excess liquidity held by commercial banks in the face of seemingly profitable investment opportunities affect the output gap’. Later on, as suggested in theory and as also empirically supported, the excess liquidity and the output gap may not necessarily be in a causal relationship, but rather in a correlation relationship. Thus, the initial research question was rephrased to ‘what is the relationship between excess liquidity and the output gap’. The critically assessed theories that investigate the behaviour of the business cycle with respect to financial factors provided the background to the conduct of the empirical analysis in this chapter. For the latter purpose, a set of three countries with macro-level data was used, including regulatory, financial, institutional and fiscal variables in the model. We endeavoured to improve on some of the weaknesses found in the empirical literature for the developed economies as well in the literature on transition economies. In particular, most of the studies conducted linking the business cycles to the financial cycles or financial variables compare the business cycle (in flows) with aggregate financial data. Based on Biggs and Mayer (2013), who argue that it is the change in credit flows i.e. - ‘credit impulse’- that should be investigated with regard to the business cycles, we addressed this issue by establishing a flow-with-flow data model. Another limitation of the empirical literature is that the majority of the studies conducted on excess liquidity, in particular the ones that include growth or output gap variables in the model, as well as those investigating the business cycle with the regard to the financial cycle do not control for contemporaneous cross-sectional correlation among the residuals. In order to consider this possible correlation among the residuals via the common observed and unobserved factors a SUR model was applied.

One of the findings of this chapter are that the regulatory and precautionary factors appear to have a more significant impact on excess liquidity in Kosovo and Estonia, while the involuntary motive prevails in the Czech Republic case. The external sector as proxied by the Eurozone output gap appear to persistently push in the same direction the business cycles in the Czech Republic and Estonia, while the domestic developments of consumption and investment as proxied by credit have the highest impact in Kosovo business cycle. Another finding of this chapter is that there is no clear transmission mechanism from excess liquidity to the output gap, thus there is no simple causal relationship between the two. The precautionary and involuntary excess liquidity and the output gap are more strongly linked via the unobserved common factors rather than by the common observables, indicating that they are located in a more complex system. The robustness of these results has been examined using different specifications, such as changing the sample size and re-estimating with SUR approach and also a simple OLS equation-by-equation approach.

Chapter 6. Conclusions and policy implications

6.1. The role and the organization of this chapter

The aim of this research project was to investigate the relationship between the output gap and excess liquidity in European transition economies. In doing so, the thesis was divided into three main parts: the first part investigated and measured potential output and the output gap, with special reference to the Czech Republic, Estonia and Kosovo. The second part defined and investigated the determinants of excess liquidity and further the analysis disaggregated between precautionary and involuntary excess liquidity. The third part was devoted to an exploration of possible theoretical and empirical links – causal, simultaneous or otherwise endogenous – between the output gap and excess liquidity. For the purpose of addressing the relationship of interest, a sequence of research questions was elaborated: are European transition economies operating with underutilised resources? Are banks in European transition countries lending at a lower rate than those in developed countries, while having partially unutilised resources, such as excess reserves? What is excess liquidity and what are its main determinants? Is unutilised excess liquidity dampening growth and increasing the output gap, or are both an outcome of a wider system?

The structure of this chapter is organized as the following. The first part presents the summary of the main findings of the thesis, followed in 6.3 by an explanation of the main contributions to the knowledge of this research programme. In the following section, the encountered limitations while conducting the research are presented, be they conceptual, theoretical, empirical or technical in nature. Additionally, the possibilities of further extending our research agenda are presented at the end of section 6.5. Finally, the chapter ends with a conclusion to the thesis.

6.2. Summary of the main findings

Given that the structure of the thesis entails three different topics brought together in a coherent sequence, the findings of the thesis can also be initially presented threefold. The first part of the findings was concerned with the results generated from what was considered the most appropriate estimation technique of the unobserved components, namely potential output and the output gap. The second part of the findings was related to the disaggregation process between precautionary and involuntary excess liquidity and hence providing an answer to whether excess liquidity is predominantly a supply or demand side problem. The third and most challenging part was concerned with defining the type of the relationship characterising the output gap and excess liquidity and based on which possible policy implications would be drawn.

Based on the arguments discussed in chapter 1, indicating that many European transition economies have been continuously operating with relatively high levels of unemployment, chronic current account deficits and relatively low levels of bank credit to GDP, we established evidence in support of our hypothesis that many of these countries may be operating with a relatively high rate of underutilised resources. In addition to the underutilised resources, market and institutional deficiencies together with frequently sluggish economic growth have also suggested that many transition economies have been persistently operating with a relatively large (negative) output gap. From here, we moved on to generating measures that indicate the general stance of economic activity, namely potential output and the output gap.

Given their unobservable nature, both potential output and the output gap had to be independently estimated by the author. Thus, the first empirical part of the research was largely concerned with finding the most appropriate estimation technique that handles the characteristics of the three sample countries' economies, namely the Czech Republic, Estonia and Kosovo. Due to an inability to cover every European transition country in the study, these three countries with different levels of economic development were selected in order to have a representative sample of the region. The Czech Republic represents the more developed EU member countries, Estonia was one of the most severely hit countries by the Global Financial Crisis and Kosovo is one of the least developed European transition countries. Given the vast number of estimation techniques available in the literature, each of which have their advantages and disadvantages, initially several univariate and multivariate measures were taken into consideration Initially the evolution of estimations of the unobserved components techniques, based on different concepts of the trend (i.e. potential output), such as linear, deterministic and stochastic trends was elaborated. Further, in order to better describe the behaviour of the observed actual output or other macroeconomic time series, other commonly used and more advanced decomposing techniques were taken into account, such are the production function, the SVAR and the Hodrick-Prescott filter and unobserved components (Kalman filtering technique) approaches. After lengthy consideration, the unobserved components operationalized via the Kalman filter was chosen as the preferred method to estimate potential output and the output gap for the three sample countries. As a widely accepted and quite powerful filtering method, this was chosen mainly because of its ability to deal with the characteristics of transition economies' data, such as possible structural breaks in the series and short time-spans because it is less data demanding and not being sensitive to the end-point estimates when data are revised.

The unobserved components model was estimated with univariate and multivariate approaches. The estimated model from the univariate model was expected to deliver relatively inferior results, compared to the multivariate model. This is because the univariate model is statistical in nature as it only exploits the real GDP series to extract the unobserved components, whereas in the multivariate model we have utilised the economic content by adding an augmented Phillips curve equation into the system of equations. This proved to be the case for the multivariate unobserved components for the Czech Republic: the estimated potential output and output gap were more informative and corresponded better to its economic situation. However, in the case of Kosovo the results did not change much in the multivariate model, suggesting that the output gap did not represent a useful indicator for inflation targeting policy-making due to the lack of the monetary policy in that country. The opposite was found for Estonia, where the current and previous period's output gap influence the level of inflation and may be important for conducting monetary policy in that country. Overall, in addition to the findings of whether inflation turned out to be relevant information when generating the output gap, we also found that, contrary to what theory suggested that the multivariate model did not always outperform the univariate specification. This is because, in addition to bringing along additional information and hence potentially additional accuracy to the estimates of potential output and the output gap, the multivariate specification also appeared to increase the complexity of the procedure and uncertainty of the findings. Nevertheless, given that it utilises additional information from the economic theory, the multivariate results were used for further analysis in the following chapters.

Estimating the output gap as a business cycle proxy in the economy was of particular importance in this research, as this enabled us to continue the investigation of whether any specific component of the underutilised resources was causing the output gap to be larger in these countries. One of the potentially most important underutilised resources is excess liquidity in the banking system of European transition economies. Having in mind that banks in these countries represent the main source of finance and that the level of credit to the private sector was found to be relatively lower than in developed countries, the holding of excess reserves was initially suspected to be associated with larger output gaps in these economies. As evidenced in chapter 1 and further elaborated in chapter 4, one of the common and continuously underutilised resources in European transition economies was the persistent accumulation of relatively high excess liquidity in the banking sector of these economies. Therefore, before exploring whether excess liquidity accumulation causes a higher output gap, the next stage of the research was initially focused on the excess liquidity phenomenon, in particular in clarifying the definition and determinants of excess liquidity in European transition economies. In this part of the research the appropriate definition of excess liquidity was reconsidered and the following research question was addressed: 'why do profit maximizing banks hold excess reserves, in the face of opportunities for seemingly profitable loans?

In trying to incorporate transition economies' features into this analysis, the supply and demand side causes of excess liquidity were reformulated: whether excess liquidity in the banking system is voluntarily held as a precautionary measure against market risk and unforeseen contingencies or involuntarily held due to a lack of demand or other factors outside commercial banks control. To account for precautionary and involuntary determinants, two separate models were estimated for each motive. The data sample containing 84 banks in 15 European transition economies were estimated via the fixed effects approach with AR(1) residuals. Even though the precautionary motive model was augmented with regulatory and risk factors, amongst other factors, the findings suggested that funding costs and growth expectations were the only significant determinants of the precautionary motive for holding excess liquidity in European transition economies. Due to the inability to obtain the discount rate or a market interest rate to approximate the banks' cost funds for all the countries in the sample, the deposit interest rate was used instead, which in transition countries can be argued to be a better proxy, given that most of banks' funds comes from domestic deposits (established in chapter 1). Conversely to the conventional wisdom suggesting that increased funding costs (as proxied by deposit interest rate) should increase precautionary excess reserves in order to avoid liquidity shortages, our findings confirmed that in European transition economies banks tend to use up their own funds and thus reduce excess reserve holdings instead of acquiring more funds (e.g. more deposits) as the opportunity cost of holding non-remunerative funds increases. The other finding suggested that growth expectations for the upcoming year play a significant role in determining the next year's expected lending and consequently determining the level of accumulated excess reserves. On the other hand, the estimated results from the involuntary motive model were in line with the theory suggesting that banks in transition economies hold involuntary excess liquidity due to involuntary accumulation of, in our case, deposits and foreign direct investments, whereas increased lending would consequently lower these involuntary funds. In order to find which of the motives prevails in European transition economies, in the second stage of estimations, the estimated models were disaggregated between precautionary and involuntary excess liquidity. The disaggregated results were generally in line with those of the Saxegaard (2006), suggesting that the involuntary motive for holding excess reserves prevails in the case of European transition economies, i.e. most of the accumulated excess liquidity was a consequence of demand-side problems. Based on the estimated results the reasons for the involuntary motive prevailing appear to be related to a lack of alternative savings vehicles, lack of absorption capacity in the economy for feasible projects, market and institutional failures, etc. The different economic, market and institutional conditions in European transition economies (and not necessarily their central banking policies) as compared to more developed European countries, may also partially explain why subsidiaries of banks with head offices in developed European countries adopted different practices in European transition economies, e.g. different level of credit and interest rates.

After exploring the definition and determinants of the output gap and excess liquidity separately, this research project went on to explore the possible relationship between the two. Due to the lack of a specific theory addressing any sort of the relationship between output gap and excess liquidity, the theoretical discussion started from a broader debate, that concerning the growth and finance nexus. During this discussion, we found that the studies examining the growth and finance relationship in a developed countries context were mainly leaning towards the conclusion that it was the financial sector that promoted growth, rather than the other way around. On the other side, studies exploring the direction of causality in developing or transition economies tended to find the opposite direction of causation; financial sector following growth prospects in the economy. A third line of argument found evidence in favour of a simultaneous relationship between growth and finance.

Trying to bring the theoretical discussion closer to our relationship of interest, an emerging theory relating business cycles with financial cycles were elaborated. These studies that emerged only after the Global Financial Crisis (GFC) criticise the mainstream theories that relate a stable economic output with stable inflation and concentrate their policy making efforts on inflation targeting. They argued that financial cycles played a crucial part in leading business cycles, including the recent financial crisis, and that were wrongly neglected in the mainstream theory and policy making. Based on these studies, the causality runs from financial cycles to business cycles. However, it was noticed that the possibility of an inverse relationship or an endogeneity issue between business cycles and financial cycles were not considered in this part of the literature. Thus, suspecting simultaneity between output gap and excess liquidity, we devoted a separate section to examining endogeneity issues between the two. Given that there is little or no theory on the relationship between output and excess liquidity we lacked theoretical guidance as to what sort of endogeneity may characterise these two variables. Further, though the endogeneity issue surrounding business and financial cycles in recent studies was not specifically addressed, Biggs and Mayer (2013) line of argument served as a starting point. Based on their arguments built mainly on descriptive statistics and comparison between stock and flow data, rather than being a causal one, it turned out that the business and financial cycles are actually highly correlated. The shift from a causal to correlation relationship changed the conceptual relationship between output gap and excess liquidity, since in this case we no longer could investigate the causal impact of excess liquidity on the output gap. This is because in a correlation relationship the causality direction may not run from one to the other or it may simply be unknown which one causes the other, because both may represent outcomes of a wider system. This correlation relationship was also one of the main findings, suggesting that for output gap and excess liquidity to change a third group of factors need to shift for them to change. Therefore, although at the beginning we had intended to investigate the impact of excess liquidity (as an unutilised resource that could be translated in more loans) on the output gap (causality running from excess liquidity to the output gap), later on, based on Biggs and Mayer's (2013) arguments and also after being able to empirically confirm (via single equation systems: FGLS and PCSE) that a causal relationship may be conceptually misconceived, the initial research question was reformulated from 'does excess liquidity impact output gap' to 'what is the relationship between the output gap and excess liquidity'?

These theoretical and empirical arguments led us to search for a new empirical modelling technique that would reflect the correlation relationship. Therefore, to take account of the correlated error terms of the equations with potentially endogenous dependent variables, this research employed a different and arguably more appropriate estimation method, that of Zellner's (1962) system-equation approach, the Seemingly Unrelated Regressions (SUR). The

SUR technique was considered to be an appropriate method in our case, as it allowed the output gap and excess liquidity to be jointly linked via common observed and unobserved determinants. Given that the SUR estimation method enjoys the virtue of estimating different slope coefficients for each cross-sectional unit, instead of treating all the data as a panel, we have employed a time-series framework for the SUR, which allowed a nine-system equation estimation of the model, three per each country: two equations standing for precautionary and involuntary excess liquidity and the third one representing an aggregate demand identity, for the output gap. In the SUR system of equations, each country was linked via the specific observables, common observed variables (credit to GDP and the GFC captured by the break and year dummies) as well as common unobserved factors as captured by the error terms.

Even though the SUR model was carried out with several different combinations of the data (shorter time span but less interpolated values or longer time span but more interpolated data, etc.), the core findings of the main specified SUR model remained as follows. The null hypothesis of zero contemporaneous covariance dependence between the errors from each equation was rejected at conventional level of significance, indicating evidence in support of contemporaneous cross-sectional correlation among the error terms and, hence, some efficiency gain from employing the SUR method. The contemporaneous cross-sectional correlation together with the year dummies, once more confirmed our initially posed hypothesis that the output gap and excess liquidity are part of a system jointly determined via underlying observed and unobserved factors.

The results from the final SUR specification indicated that stronger regulatory actions like increasing the reserve requirement ratio and equity to assets tend to reduce excess liquidity growth, as they themselves represent additional precautionary buffers as well as add to the banks' (opportunity) costs. Higher funding costs, as approximated by the deposits interest rate, contrary to what theory suggested, appeared to reduce the precautionary excess liquidity in all three countries, a finding consistent with the results in chapter 4. The risk indicators proxied by volatility measures appeared to have an insignificant impact in all three sample countries, possibly due to only a small effect being captured by the data or simply because these were poor proxies for our sample of countries, despite being suggested by orthodox theory. The non-performing loans, as a typical risk measure for banking sector, appeared to have a significant impact on precautionary excess liquidity in Estonia and Kosovo, though the direction of the impact differed. The results for involuntary excess liquidity turned out to be less informative, given the statistical insignificance of most of the variables and the counterintuitive signs in the case of credit and deposits coefficients.

Overall, the findings of this chapter are that the regulatory and precautionary factors appear to have a more significant impact on excess liquidity in Kosovo and Estonia, while the involuntary motive prevails in the Czech Republic case. Regarding the output gap equation, the estimated results consistently indicated (in the main specified model as well as in robustness checks) that the business cycles of Czech Republic and Estonia are strongly driven by the external environment events (Eurozone business cycle), whereas domestic developments (consumption and investments) were the main driver in Kosovo.

In addition to the common and specific observable factors, we found that the unobserved factors driving output gap and excess liquidity jointly as a system were even more pronounced than the observed part of the model. A careful analysis of the correlations in the Breusch-Pagan test, as well as information picked up by the year dummies (which do not encompass information on any specific policy variable), all pointed out that output gap and excess liquidity were jointly affected by the unobserved part of the model, be it in the same or opposite direction.

6.3. Main contributions to knowledge

This research project makes conceptual, theoretical as well as empirical contributions to knowledge; these will be separately elaborated in the following sub-sections.

6.3.1. Conceptual contributions to knowledge

The first conceptual contribution of this research project consisted in approaching the business cycle concept from a different direction, which contrasts with the approach of mainstream theory. Mainstream theory has usually related the importance of potential output and the output gap with the conduct of monetary and fiscal policy. Though we have acknowledged their ultimate importance in overall macroeconomic decision-making and based our empirical framework initially on mainstream theory, the primary reason for estimating potential output and the output gap was to establish evidence of the extent of below potential economic activity in European transition economies.

The second contribution to knowledge consists in redefining the concept of excess liquidity as well as disaggregating the precautionary and involuntary motive in the European transition region. Previous studies had not paid much attention to the definition of excess liquidity, indeed the excess liquidity phenomenon itself has been rarely examined and when treated it has mainly been discussed in the light of monetary policy effectiveness. We found that the definition of excess liquidity was not a straight forward matter. The limited literature offered two different definitions: the first classifying as excess reserves the part of above statutory deposits held at the central bank and the defining excess liquidity as total commercial bank liquid assets minus the required liquid assets determined by the central bank. The second definition, in contrast, implied the inclusion of other parts of liquid assets, such as treasury bills which can be quickly converted into commercial banks' reserves or loans. Criticising both of these definitions as imprecise, a contribution to knowledge of this research was in clarifying the concept of excess liquidity and subsequently redefining its measurement. Based on our proposed definition, excess liquidity includes the above statutory requirements of deposits at central banks (total reserves held at the central bank minus reserve requirement ratio) plus the part of liquid assets held by the commercial banks (cash at commercial banks' accounts, treasury bills, securities, derivatives and placements abroad) minus the mandatory liquidity ratio set by the regulator or other restrictions. This new definition subsequently affected the measurement of excess liquidity, which is different to the other measurement approaches in the literature.

6.3.2. Theoretical contributions to knowledge

The first theoretical contribution to knowledge consists in relating the discussion of excess liquidity with the mainstream theories. While critically assessing the theory underpinning the determinants of excess liquidity, we noticed that other studies had raised fundamental questions, such as 'is excess liquidity a supply or a demand side problem' without citing the mainstream macroeconomic theories from where these questions emanate. Therefore, in the theoretical discussion we added to the current literature by initially presenting the discussion from (New) Keynesian and (New) Classical perspectives. We argued that the Classical position, serving as an explanation based upon supply side factors, suggesting that there would be no involuntary excess liquidity accumulated in the commercial banks, just a state of equilibrium at the prevailing interest rate. In this case it would be useless for the central banks to intervene to reduce the excess liquidity below what may be an equilibrium level. Thus, solutions should be found in the capital markets and banking systems respectively, i.e. to strengthening competition and eliminating central bank restrictions. Based on the Keynesianism view, excess liquidity would be regarded from the level of aggregate demand that creates equilibrium between investment and saving plans. Based on this theory, excess liquidity could be viewed either as a temporary deviation from the equilibrium, or as a consequence of a systemic market failure to adjust savings and investments, which then would manifest in a build-up of excess liquidity. The savings of the public in the banking system and the loans from the banks may be considered to be in equilibrium at any level of excess liquidity accumulated; however this could create an equilibrium in which not all resources are utilised.

The second theoretical contribution to knowledge consists in understanding the conceptual and theoretical relationship between output gap and excess liquidity. This is because, to the best of our knowledge, this specific relationship had not been previously investigated. The lack of specific theory together with a lack of guidance regarding the appropriateness of an empirical approach added to the difficulty of examining our relationship of interest. While defining the sort of the relationship between our variables of interest we have also added two novelties to the literature. Firstly, hypothesising a new conceptually endogenous relationship between the output gap and excess liquidity and secondly, in order to account for potential endogeneity between the two suggesting an alternative econometric method not previously used in the context of business cycle analysis. Given that excess liquidity, if translated into further loans for the economy is believed to impact on growth and production capacity, initially a possible causal relationship from excess liquidity to the output gap was investigated, followed by a more thorough analysis of a potential system implications between the two. Thus, when exploring the relationship between output gap and excess liquidity, the latter was treated as potential (unutilised) additional credit for the private sector of the real economy, which to the best of our knowledge also represents a new perspective. Initially, believing that the output gap and excess liquidity are in a causal relationship, we tried out several single-equation modelling approaches. Suspecting simultaneity between the two variables of interest, we found out that just as much as excess liquidity causes the output gap, the output gap also causes excess liquidity. Based on our empirical findings and also mainly relying on the Biggs and Mayer (2013) arguments, we found out that treating output gap and excess liquidity form a single causal point of view may be conceptually misconceiving. Pursuing a single-equation approach was therefore also inappropriate as this approach could no longer take into account the endogenous relationship between output gap and excess liquidity. From this point the initially posed research question viewing two variables of interest ‘does excess liquidity cause higher output gap’ evolved into the question 'what is the relationship between output gap and excess liquidity'. Biggs and Mayer (2013) had argued that the business cycle and credit impulse (new loans) are closely correlated to each other, to the point where one can approximate the other. Therefore, the ultimate contribution to knowledge of this research consists in providing a different conceptual relationship and rather more complex than initially thought with regard to output gap and excess liquidity. This is because, a correlation relationship indicates that both variables are an outcome of a wider system and do not necessarily cause each other.

6.3.3. Empirical contributions to knowledge

The first empirical contribution to the knowledge consists in implementing the Kalman filtering method to examine the business cycle in Kosovo, which was novel for this country. To the best of our knowledge, this is also the first attempt to estimate the business cycle in Kosovo. The estimated results suggested that from 2002 to 2008 the actual output was performing below its potential in Kosovo; around 2009 the economy was almost reaching its equilibrium followed by a short period of overheating; and soon after in 2012 the economy went back to below potential activity. The generated results for the Czech Republic and Estonia indicated a rather different behaviour of the business cycle: during the 2000s these two countries were enjoying robust economic growth and above potential economic activity up to the onset of the Global Financial Crisis (GFC), the latter produced a turning point in the business cycle and a period of below potential economic activity.

The second empirical contribution to knowledge is estimating the determinants of excess liquidity in European transition economies, a sample of countries different from previous studies as well the use of a different data, source: Bankscope. The estimated results indicated that regulatory variables like reserve requirement ratio and growth prospects are two main factors inducing banks in European transition economies to hold precautionary excess reserves, whereas deposits, lending interest rates and FDIs are the main factor leading banks to accumulate excess reserves involuntarily. Furthermore, our findings suggested that involuntary motive prevails in inducing banks in these countries to accumulate excess liquidity.

The third empirical contribution to knowledge consists in finding the most appropriate method to estimate the model that would reflect the endogenous-correlation relationship between output gap and excess liquidity. Even though Biggs and Mayer (2013) had offered a different perspective of what the relationship between business cycles and financial (credit) cycles may be, they reached this conclusion based purely on descriptive statistics. After a due consideration, in the current research a system-equation approach, Seemingly Unrelated Regressions (SUR) method, which accounts for endogeneity between output gap and excess liquidity, was considered to be the most appropriate method to estimate the model. This approach did not provide a causal mechanism, however allowed equations of output gap and excess liquidity to be correlated via the common observed and unobserved determinants. The use of the SUR technique represents another contribution to the knowledge, because to the best of our knowledge it has not been previously used in this field of investigation, even in empirical studies of developed economies.

6.4. Policy implications

The first policy implication of this research is related to the first part of the thesis: that the output gap in Kosovo case does not appear to impact on the inflationary pressures in the economy. The insignificant relationship between output gap and the inflation rate may also reflect the specific labour market in Kosovo. Even though the unemployment rate is very high, the reservation wage is also thought to be relatively high due to the effect of remittances, whereas wages are not necessarily flexible in the short-run. Therefore, the impact of the output gap on the inflation rate may be insignificant. Therefore, an inflation targeting policy in Kosovo is not feasible, because inflationary pressures will be mainly determined by external factors.

The second policy implication generated from this research is related to the status of excess liquidity. Based on our findings, the prevailing reason why commercial banks in European transition economies accumulate excess liquidity is of an involuntary nature, i.e. a demand side problem. Given that the causes of accumulating involuntary excess liquidity are mainly found outside of banks' control (lack of demand for loans, lack of alternative vehicles to invest, deficient markets and institutions, etc.), policies such as lowering the reserve requirement or discount window may not necessarily be feasible. This is because as the causes for involuntary excess liquidity lie outside the scope of central banks' policymaking or commercial banks' will to reduce excess liquidity and extend lending. Instead the accumulation of excess liquidity by commercial banks in European transition economies may be better addressed through improving their legislative framework, which in turn may increase the efficiency of the intermediation process. In addition, providing a better economic environment for investment may also help to stimulate the demand for loans.

The distinction between precautionary and involuntary excess liquidity has important policy implications. The precautionary portion of excess liquidity, even though fairly small in our sample of European transition economies, may entail an inefficient allocation of resources. This is because, if conditions that give rise to banks' uncertainties improve, this part of excess liquidity could be quickly mobilised for further lending. If banks hold excess reserves only for precautionary purposes, then a loosening of monetary policy (e.g. lowering the reserve requirement) may be effective, since it would increase excess liquidity above the level demanded by commercial banks for precautionary purposes. In this case, banks are expected to expand lending by lowering the cost of borrowing or reducing the rationing of loans. Similarly, a contractionary monetary policy would lead banks to contract lending to maintain their desired level of excess reserves.

On the other hand, if the holdings of excess liquidity are involuntary, as is predominantly the case here, in the sense that banks are unable to expand lending, then attempts by central banks to boost credit demand by lowering the cost of borrowing will be largely ineffective. An expansionary monetary policy in that case would simply inflate the level of unwanted excess reserves in commercial banks and not lead to an expansion of lending. In this case the best policy response is, as Heenan (2005) points out, to address the underlying causes of involuntary excess liquidity policies. Such policies could entail improving the efficiency and the information structure of the commercial banking system by increasing competition in the financial sector or addressing the lack of well-developed interbank market (Saxegaard, 2006). In that case a likely outcome may be a higher lending rate due to increased demand for loans, thus raising the opportunity cost of excess liquidity. Otherwise, an improvement in the economic outlook reduces the riskiness of new borrowers.

The principal finding of this research project is that rather than being in a causal relationship, where less excess liquidity would be reflected in a lower output gap, these variables are part of a system associated through underlying observed and unobserved factors. Therefore, the third and most important policy implication of this research is that the relationship between output gap and excess liquidity is not causal, i.e. reducing excess liquidity will not necessarily lead to a smaller (negative) output gap. This is because, the output gap and excess liquidity are bound together in a correlation relationship, which means that both are an outcome of the economic system or both reflect the economic system. If the relationship between output gap and excess liquidity is one of correlation, then instead of pushing banks to reduce excess liquidity and increase lending so the output gap may decrease, other ‘third party’ factors need to change so that our two variables of interest will improve. However, any potentially functional policy variable that we could suggest for policymakers was captured by the unobserved part of the model and thus not identified and consequently not yet very well understood. Thus, there seems to be no straightforward policy framework informed by a clear transmission mechanism from one to the other. Therefore, a policy whereby the regulatory authority induces banks to reduce excess liquidity in order to increase lending and thus decrease the output gap may not have the desired effect or may even produce quite unintended consequences, such as causing inefficient lending e.g. a new wave of bad loans or causing a higher (negative) output gap. Instead, a more complex policy framework is needed where other factors outside of the system need to be identified that are capable of pushing both variables in the desired direction. For this reason, we need a better theoretical understanding and additional empirical findings to properly understand the underlying relationship between output gap and excess liquidity. Hence, the avoidance of a policy where banks are pushed to reduce excess liquidity is the main policy implication of this chapter, since that may lead into a new wave of bad loans rather than impetus to economic activity.

6.5. Limitations and future work

The theoretical and methodological limitations that may have influenced the interpretation of our findings are reported in this section. First, as noted in chapter 3, when estimating unobserved components, which in our case were potential output and the output gap, no matter which method we would use or how complex and sophisticated the decomposition technique would be, each method entails its own advantages and disadvantages. Furthermore, given that they represent estimates and not the actual measures, any generated series of potential output and the output gap are at best uncertain (Anderton et al., 2014), therefore we cannot claim that the generated series accurately represent the true state of economy. The estimates of potential output and the output gap in Kosovo case may be prone to additional measurement error, given the lack of quarterly data. The Kosovo Statistical Agency started publishing quarterly GDP data only at the end of 2014 for a few recent years. The disaggregated quarterly data from annual data used for Kosovo case were estimates from the CBK staff.

Second, the generated results from chapter 4 on the determinants of excess liquidity utilised data from the Bankscope database, which is known for the disadvantage of having a considerable amount of missing values. The financial reports of the banks reported from transition economies provide an even more severe case of missing values. Thus, having in mind that we have used a dataset of 39 banks in 15 European transition economies, the large number of missing values led to a relatively small dataset.

Third, in chapter 5, when constructing the model that relates output gap with excess liquidity, some of the data especially at the beginning of the series and particularly in Kosovo's case were missing. Because we employed the SUR modelling approach, which requires balanced data for the covariance matrix elements to parameterise, the missing values have limited us in using a relatively short time-span and therefore a smaller dataset. In an attempt to utilise the largest dataset possible with a longer time-span and to not remove around one third of the data because of the missing values, most of the data at the beginning of period (2004 – 2006) for Kosovo and a very few ones for the Czech Republic and Estonia had to be interpolated or imputed when missing. Even though data imputation and interpolation are widely used methods by researchers when similar issues arise, and even more so in transition context (Ilzetski et al., 2013), they represent estimates and not actual values, thus add to the measurement error of our estimates.

Fourth, the estimated results in chapter 5 with SUR modelling approach indicated that the nine equations in the model had a significantly high explanatory power, nevertheless the individual variables in the nine equations were rarely individually significant. Adding here the counter intuitive coefficient signs of some of the variables in the model, led us to believe that multicollinearity may have been a problem of our data. However, multicollinearity is an issue that cannot be addressed in small datasets.

Overall, rather than providing answers on the way forward for policymakers, this research leaves as open-ended way whether pushing banks to lend more would do much good for the economy. Without intending to over-claim the findings of this research, our research puts a question mark over any work previously done regarding the relationship between business and financial cycles and specifically regarding the relation between output gap and excess liquidity, particularly in the context of the aftermath of the GFC. Indeed, the whole macroeconomic system needs rethinking, even beyond the transition context, especially with regard to the aftermath of the financial crisis, since this system is not yet very well understood. Therefore, the current state of both theory and empirical evidence on the relationship between the output gap and excess liquidity still do not provide a sufficient basis for forming policy guidance. This is because previous studies are scarce and our estimated results give directions towards other unobserved determinants that push both, the output gap and excess liquidity in the same or opposite direction. Consequently, since these 'third factors' were not identified from the observed but rather unobserved part of the model, before providing any firm direction for policymaking a much better theoretical framework is needed. However, the exploratory journey undertaken in this research represents only the beginning of this research agenda.

Finally, it should be noted that after exploring the relationship between output gap and excess liquidity, another chapter of the thesis was originally planned as the last part of the research. The aim of that chapter was to investigate the capacity utilisation of the firms in European transition economies, by employing the World Bank's BEEPS (2002 – 2009) dataset. After investigating the business cycle in the economy at the macroeconomic level, an extended investigation on the determinants of capacity utilisation and capacity utilisation growth would add a firm level dimension to the research. However, due to the demands of answering the questions set for the other three parts of the planned research and the overall wordage limitation on a PhD thesis this part of the research could not be undertaken. The determinants of the capacity utilisation of the firms in European transition economies are planned to be part of the further research in the post-doctoral studies phase.

References

Acock, A. (2013) *Discovering Structural Equation Modeling using Stata*, Texas: Stata Press.

Agenor, P., Aizenman, J. and Hoffmaister, A. (2004) The Credit Crunch in East Asia: What Can Bank Excess Liquid Assets Tell Us? *Journal of International Money and Finance* 1(23) pp. 27–49.

Aguiar, M. and Gopinath, G. (2007) The Role of Interest Rates and Productivity Shocks in Emerging Market Fluctuations *Central Bank of Chile Working Papers* No. 445.

Aikaeli, J. (2006) Determinants of Excess Liquidity in Tanzanian Commercial Banks *University of Dar es Salaam* (mimeo)

Akram, M., Hussain, S., Raza, S. and Masood, S. (2014) An Empirical Estimation of Okun's Law in Context of Pakistan *Journal of Finance and Economics*, 2(5) pp. 173 – 177.

Alichi, A. (2015) A New Methodology for Estimating the Output Gap in the United States *IMF Working Paper* No. 144.

Anderton, R., Aranki, T., Dieppe, A., Elding, C. Haroutunian, S., Jacquinot, P., Jarvis, V., Labhard, V., Rusinova, D. and Szörfi, B. (2014) Potential Output from a Euroarea Perspective *European Central Bank Occasional Paper Series* No. 156.

Antonicova, Z. and Hucek, J. (2005) Multivariate Filter with Unobserved Components *BIATEC* 6(13) pp. 13 – 18.

Apel, M. and Jansson, P. (1999) A Theory-Consistent System Approach for Estimating Potential Output and the NAIRU *Economics Letters* 3(64) pp. 271-275.

Ball, L., Leigh, D. and Loungani, P. (2013) Okun's Law: Fit at 50? *IMF Working Paper* No. 10.

Baltagi, B. and Li, Q. (1991) A Joint Test for Serial Correlation and Random Individual Effects *Statistics and Probability Letters* 3(11) pp. 277-280.

Baltagi, B. (2008) Econometric Analysis of Panel Data *Business and Economics* Hoboken: John Wiley and Sons.

Bank of Albania (2015) Monetary and Financial Statistics, accessed in April 2015 http://www.bankofalbania.org/web/Time_series_22_2.php?evn=agregate_parent_sel&evb=agregate&Cgroups=36&periudha_id=5

Bank of Serbia (2015) Monetary Sector Statistics, accessed in April 2015

Bank of Lithuania (2015) Monetary Financial Institutions Balance Sheet and Monetary Statistics, accessed in April 2015 http://www.lb.lt/monetary_financial_institutions_balance_sheet_and_monetary_statistics

Basarac, M., Skrabic, B. and Soric, P. (2011) The Hybrid Phillips Curve: Empirical Evidence from Transition Economies *Czech Journal of Economics and Finance* 4(61) No. 367.

Battaglini, M. and Coate, S. (2008) Fiscal Policy over the Real Business Cycle: A Positive Theory *NBER Working Paper* No. 14047.

Baum, C. (2001) Residual Diagnostics for Cross-Section Time Series Regression Models *The Stata Journal* 1(1) pp. 101 – 104.

Baum, C. (2006) An Introduction to Modern Econometrics Using Stata *Stata Press* Texas: StataCorp.

BBC (2014) Budget 2014: Output Gap Judging a Nation's Economy *BBC*, March 2014, http://www.bbc.com/news/uk-politics-26613739

Beck, N. and Katz, J. (1995) What to Do (and Not to Do) with Time-Series Cross-Section Data *The American Political Science Review* 3(89) pp. 634 – 647.

Beck, T. (2012) Finance and growth – lessons from the literature and the recent crisis *LSE Growth Commission* (mimeo).

Beck, T., Demirgüç-Kunt, A. and Maksimovic, V. (2005) Financial and Legal Constraints to Firm Growth: Does Firm Size Matter? *Journal of Finance* 1(60) pp. 137 – 177.

Beck,T. and Demirguc-Kunt, A. (2006) Small and Medium-Size Enterprises: Access to Finance as a Growth Constraint *Journal of Banking and Finance* 11(30) pp. 2931 – 2943.

Bencivenga, V. and Smith, B. (1991) Financial Intermediation and Endogenous Growth *Review of Economic Studies* 2(58) pp. 195 – 209.

Benes, J. and Ndiaye, P. (2004) A Multivariate Filter for Measuring Potential Output and the NAIRU: Application to The Czech Republic *IMF Working Paper* No. 45.

Benes, J., Clinton, K., Garcia-Saltos, R., Johnson, M., Laxton, D., Manchev, P. and Matheson, T. (2010) Estimating Potential Output with a Multivariate Filter *IMF Working Paper* No. 285.

Berger, A. and DeYoung, R. (1997) Problem Loans and Cost Efficiency in Commercial Banks *Journal of Banking and Finance* 6(21) pp. 849 – 870.

Berger, A., Earle, J. and Sabirianova, K. (2001) Worker Training in a Restructuring Economy: Evidence from the Russian Transition *IZA Discussion Paper Series* No. 361.

Berglof, E. and Bolton, P., (2002) The Great Divide and Beyond: Financial Architecture in Transition *Journal of Economic Perspectives* 1(16) pp. 77-100.

Bernanke, B., Gertler, M. and Gilchrist , S. (1998) The Financial Accelerator in a Quantitative Business Cycle Framework *NBER Working Paper* No. 6455.

Bernanke, B., Gertler, M. and Gilchrist, S. (1999) The Financial Accelerator in a Quantitative Business Cycle Framework in Tayler, J. and Woodford, M. (eds.) *Handbook of Macroeconomics 1,* Amsterdam: Elsveier Science.

Beveridge, S. and Nelson, C. (1981) A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle *Journal of Monetary Economics* 7 pp. 151 – 74.

Bevilacqua, F. and Zon, A. (2001) Random Walks and Non-Linear Paths in Macroeconomic Time Series: Some Evidence and Implications *Maastricht Economic Research Institute on Innovation and Technology* (mimeo).

Bezovska, A., Bogoev, J., Mitreska, A. and Vojnovik, M. (2011) Investigating the Cyclical Behavior of Fiscal Policy in the Republic of Macedonia during the Period of Transition *Croatian Economic Survey* 1(13) pp. 47 – 104.

Bhattarai, K. (2011) Panel Data Models *Econometric Analysis, Hull University, March (mimeo.)*

Biggs, M. and Mayer, T. (2010) The Output Gap Conundrum *OECD Working Paper* No. 699.

Biggs, M. and Mayer, T. (2013) Bring credit back into the monetary policy framework! *Political Economy of Financial Markets policy Brief,* Political Economy of Financial Markets Programme, St. Antony's College, Oxford University.
Billmeier, A. (2004) Ghostbusting: Which Output Gap Measure Really Matters? *IMF Working Paper* No. 146.

Blackwell, L. (2005) Estimation and Testing of Fixed-Effect Panel-Data Systems *Stata Journal* 2(5) pp. 202 – 207.

Bokan and Ravnik (2012) Estimating Potential Output in the Republic of Croatia Using a Multivariate Filter *Croatian National Bank Working Papers* No. 35.

Bolland, P. and Connor, B. (1997)A Constrained Network Neural Network Kalman Filter for Price Estimation in High Frequency Financial Data International *Journal of Neural Systems* 4(8) pp. 399 – 415.

Bonin, J. (2001) Financial Intermediation in Southeast Europe: Banking on the Balkans *The wiiw Balkan Observatory Working Papers* No. 006.

Bonin, J., and Wachtel, P., (2003) Financial Sector Development in Transition Economies: Lessons from the First Decade. *Financial Markets, Institutions and Instruments, Vol. 12, pp. 1-43.*

Borio, C. (2012) The Financial Cycle and Macroeconomics: What Have we Learnt? *Bank for International Settlements Papers* No. 395.

Borio, C., Disyatat, P. and Juselius, M. (2013) Rethinking Potential Output: Embedding Information About the Financial Cycle *Bank for International Settlements Working Papers* No. 404.

Born, B. and Pfeifer, J. (2014) Policy Risk and the Business Cycle *Journal of Monetary Economics* 68 pp. 68 – 85.

Bouis, R., Cournede, B. and Chriestensen, K. (2012) Implications of Output Gap Uncertainty in Times of Crisis *OECD Working Paper* No. 977.

Boulila, G. and Trabelsi, M. (2004) The Causality Issue in the Finance and Growth Nexus: Empirical Evidence from Middle East and North African Countries *Review of Middle East Economics and Finance* 2(2).

Box, E. and Jenkins, G. (1976) Time Series Analysis: Forecasting and Control, *Revised Edition*, San Francisco: Holden Day, U.S.A.

Bratkowski, A., Grosfeld, I. and Rostwoski, I. (2000) Investment and Finance in de Novo Private Firms: Empirical Evidence from the Czech Republic, Hungary and Poland *Economics of Transition* 1(8) pp.101 – 116.

Buddelmeyer, H., Jensen, P. Oguzoglu, U. and Webster, E. (2008) Fixed Effects Bias in Panel Data Estimators *IZA Discussion Paper Series* No. 3487.

Cameron, S. (2005) *Econometrics* Maidenhead: McGraw-Hill Education, UK.

Bureau Van Dyjk (2014) Bankscope database, accessed August 2014 https://bankscope.bvdinfo.com/version-2015622/home.serv?product=scope2006

Canova, F. (1998) Detrending and Business Cycle Facts *Journal of Monetary Economics* 41(3) pp. 513 – 532.

Caviglia, G., Krause, G., and Thimann, C., (2002) Key Features of the Financial Sectors in EU Accession Countries *European Central Bank Financial Sectors in EU Accession Countries* No. 1 pp.15 – 30.

Cevik, E., Dibooglu, S. and Barisik, S. (2013) Asymmetry in the Unemployment–Output Relationship Over the Business Cycle: Evidence from Transition Economies *Comparative Economic Studies* pp. 1 – 25.

Chen, X. and Mills, T. (2012) Measuring the Euro Area Output Gap using Multivariate Unobserved Components Models Containing Phase Shifts *Empirical Economics* 2(43) pp. 671-692.

Claessens, S., Demirgüç-Kunt, A., and Huizinga, H., (2001) How does Foreign Entry Affect Domestic Banking Markets? *Journal of Banking and Finance, Vol. 25: 891-991.*

Clark, P. (1987) The Cyclical Component of U.S. Economic Activity *Quarterly Journal of Economics* 4(102) pp. 797 - 814.

Clemente, J., Montanes, A., and Reyes, M., (1998) Testing for a Unit Root in Variables

with a Double Change in the Mean *Economics Letters* 59 pp. 175-182.
Clark, P. (1989) Trend Reversion in Real Output and Unemployment *Journal of Econometrics* 40 pp. 15 - 32

Cochrane, J. (1988) How big is the random walk in GDP? *Journal of Political Economy* 5(96) pp. 893 - 920.

Cochrane, J. (1991) A Critique of the Application of Unit Root Tests *Journal of Economic Dynamics and Control* 15 pp. 275 – 284.

Cochrane, J. (2005) Time Series for Macroeconomics and Finance *Graduate School of Business University of Chicago* (mimeo).

Cochrane, J. (2012) A Brief Parable of Over-Differencing *University of Chicago Booth School of Business and NBER* (mimeo).

Commandeur, J.and Koopman, S. (2007) *An Introduction to State Space Time Series Analysis* Oxford:Oxford University Press, UK.

Cottarelli, C., Dell'Ariccia, G. and Vladkova-Hollar, I. (2005) Early Birds, Late Risers, and Sleeping Beauties: Bank Credit Growth to the Private Sector in Central and Eastern Europe and the Balkans *Journal of Banking and Finance* 29(1) pp. 83 – 104.

Czech National Bank (2015) ARAD Time Series Database, accessed in April 2015
http://www.cnb.cz/cnb/STAT.ARADY_PKG.STROM_KOREN?p_lang=EN

Darvas, Z. and Vadas, V. (2003) Univariate Potential Output Estimations for Hungary *Magyar Nemzeti Bank Working Paper* No. 8.

De Brouwer, G. (1998) Estimating Output Gaps *Reserve Bank of Australia Discussion Paper* No. 9809.

De Masi, P. (1997) IMF's Estimates of Potential Output: Theory and Practice *IMF Working Paper* No. 177

Demetriades, P. and Hussein, K. (1996) Does Financial Development Cause Economic Growth? Time-Series Evidence from 16 countries *Journal of Development Economics* 2(51) pp. 387 -411.

Deutsche Bank Research (2011) Emerging markets banks: Strong Growth Potential Econometric Analysis *International Topics*, *Current Issues* Frankfurt, Deutsche Bank.

De Vroey, M. and Malgrange, P. (2011). The History of Macroeconomics from Keynes's General Theory to the Present *UCL Discussion Paper* pp. 2011-28.

Domaç, I., Ferri, G. and Wei, D. (1998) Is There a Credit Crunch in East Asia *World Bank Policy Discussion Paper* No. 1959.

Dornbusch, R. and Fischer, S. (1993) Moderate Inflation *World Bank Economic Review* 1(7) pp. 1-44.

European Bank for Reconstruction and Development (2005) *Annual Report 2005*, London: EBRD, UK.

ECB (2015) Statistical Bulletins 2000 – 2014, accessed in April 2015 http://sdw.ecb.europa.eu/reports.do?node=1000004045

Economic Policy Committee (2001) Report on Potential Output and the Output Gap, ECFIN/EPC/670/01/en.
http://ec.europa.eu/economy_finance/publications/publication6586_en.pdf

Efendic, A., Pugh, G. and Adnett, N. (2011) Confidence in Formal Institutions and Reliance on Informal Institutions in Bosnia and Herzegovina An Empirical Investigation Using Survey Data *Economics of Transition* pp. 1 – 20.

European Commission (2015) EU Candidate and Potential Candidate Countries' Economic Quarterly *Economic and Financial Affairs* 2015/1, Brussels: EC.

Eesti Pank (2015) Financial Sector Statistics, accessed in April 2015 http://statistika.eestipank.ee/?lng=en#treeMenu/FINANTSSEKTOR

Eggertsson, G. and Woodford, M. (2003) The Zero Bound on Interest Rates and Optimal Monetary Policy *Brookings Papers on Economic Activity* 1(34) pp. 139 - 235.

El-Ganainy, A. and Weber, A. (2010) Estimates of the Output Gap in Armenia with Applications to Monetary and Fiscal Policy *IMF Working Paper* No. 197.

Engle, R. and Watson, M. (1981) A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates *Journal of the American Statistical Association* 376(76) pp. 774 – 781.

Estrin, S., Hanousek, J., Kočenda, E. and Svejnar, J. (2009) The Effects of Privatization and Ownership in Transition Economies *Journal of Economic Literature* 3(47) pp. 699 – 728.

European Bank for Reconstruction and Development (2015) BEEPS Panel Questionnaire 2009, accessed in September 2014, http://ebrd-beeps.com/panel-data/

European Central Bank (2014) *Monthly Bulletin January 2014*, Frankfurt: ECB.

EUROSTAT (2015) Quarterly National Accounts, accessed in April 2015 http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=namq_10_gdp&lang=en

EViews 7 (2010) User's Guide I *Quantitative Micro Software,* Irvine CA: IHS Global Inc.

E-Views 8 Estimation, Forecasting, Statistical Analysis, Graphics, Data Management, Simulation *Getting Started*, Irvine CA: IHS Global Inc..

Falcetti, E., Lysenko, T. and Sanfey, P. (2005) Reforms and Growth in Transition: Re-Examining the Evidence *EBRD Working Paper* No. 90.

Federico, P., Carlos V. and, Vuletin, G. (2014) Reserve Requirement Policy over the Business Cycle *NBER Working Paper* No. 20612.

Filer, R., Gylfason, T, Jurajda, Š. and Mitchell, J., (2001) Markets and Growth in the Post-Communist *World Center for Economic Research & Graduate* (mimeo).

Fioramanti and Vicarelli (2011) The New Stability and Growth Pact: Primum Non Nocere *CEPS Working Document* No. 344.

Fischer, S., Sahay, R. and Vegh, C.(1996) Economies in Transition: The Beginnings of Growth *American Economic Review* 2(86) pp. 229 – 233.

FRED (2013) Economic Data, accessed in November 2013 https://research.stlouisfed.org/fred2/

Fuentes, R., F. Gredig, and M. Larrain (2007) Estimating the Output Gap for Chile Central *Bank of Chile Working Paper* No. 455.

Furlanetto, F., Gelain, P. and Sanjani, T. (2014) Output Gap in Presence of Financial Frictions and Monetary Policy Trade-offs *IMF Working Paper* No. 128.

Gerlach, S. and F. Smets (1997) Output Gaps and Inflation: Unobservable-Components Estimates for the G-7 countries *Bank for International Settlements* (Draft version).

Giannini, M. and Vitali, L. (2014) Productivity Growth of Transition Economies: An Assessment *Advances in Economics and Business* 3(2) pp. 133 – 147.

Ghosh, S. and Ghosh, A. (2000) East Asia in the Aftermath: Was There a Crunch? *Deutsche Bank Research Notes* No. 00-5.

Greene, W. (2008) *Econometric analysis*, New Jersey: Prentice Hall.

Grewal, M. and Andrews, A. (2011) Kalman Filtering *Theory and Practice Using MATLAB*, Hoboken: Wiley-IEEE Press.

Giorno, V., Richardson, P., Roseveare, D. and Noord, P. (1995) Potential Output, Output Gaps and Structural Budget Balances *OECD Economic Studies* 24(1) pp. 167 - 209.

Gordo, R. (1997) The Time-Varying NAIRU and Its Implications for Economic Policy *The Journal of Economic Perspectives*, 1(11) pp. 11-32

Gorton, G. and Winton, A. (1998) Banking in Transition Economies: Does Efficiency Require Instability? *Journal of Money, Credit, and Banking* 30(3) pp. 621 – 650.

Graff, M. and Sturm, J. (2010) The Information Content of Capacity Utilisation Rates for Output Gap Estimates *KOF Working Papers* No. 269.

Gréboval, D. and Munro, G. (1999) Overcapitalization and Excess Capacity in World Fisheries: Underlying Economics and Methods of Control, in: Gréboval, D. (eds.): *Managing Fishing Capacity: Selected Papers on Underlying Concepts and Issues*, Rome: FAO Fisheries Technical Paper.

Greenwood, J. and Jovanovic, B. (1990) Financial Development, Growth, and the Distribution of Income *Journal of Political Economy* 98(5) pp. 1076 - 1107.

Greenwood, J. and Smith, B. (1997) Financial Markets in Development, and the Development of Financial Markets *Journal of Economic Dynamics and Control* 21 pp. 145 – 181.

Grewal, M. and Andrews, A. (2011) Kalman Filtering: Theory and Practice Using MATLAB *Technology & Engineering* Hoboken: John Wiley and Sons.

Hamilton, J. (1994a) State Space Models, in Engle, in R. and McFadden, M. (eds.) *Time Series Analysis,* Princeton; Princeton University Press, Elsevier Science.

Hamilton, J. (1994b) *Time Series Analysis* Princeton: Princeton University Press.

Harvey, A. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter* Cambridge: Cambridge University Press.

Harvey, A. and Jaeger, A. (1993) Detrending, Stylized Facts and the Business Cycle *Journal of Applied Econometrics* 3(8) pp. 231 - 247.

Harvey, A. and Proietti, T. (2005) Reading on Unobserved Components Models *Advanced Texts in Econometrics* Oxford: Oxford University Press.

Harvey, A. and Todd, P. (1983) Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study *Journal of Business and Economic Statistics* 4(1) pp. 299 -307.

Harvey, A., Trimbur, T. and Dijk, H. (2007) Trends and Cycles in Economic Time Series: A Bayesian Approach *Econometric Institute Research Papers* No. 27.

Hashi, I. and Toci, V. (2010) Financing Constraints, Credit Rationing and Financing Obstacles: Evidence from Firm-Level Data in South-Eastern Europe *Journal of Economic and Business Review* 1(12) pp. 29 – 60.

Havolli, S. and Kastrati, A. (2012) The Capacity Utilization of the Producing Firms in Western Balkans: BEEPS Results 2009, *Balance of Payments Report No. 1*, Prishtina: CBK.

Heijdra, B. (2009) Real Business Cycles *Foundations of Modern Macroeconomics* Oxford: Oxford University Press.

Heenan, G. (2005) Dealing with Excess Liquidity, *Presentation to the Joint Africa Institute, September (mimeo).*

Hellström (2009) Financial accounting quality in a European transition economy: the case of the Czech Republic *Stockholm School of Economics, Centre for Financial Analysis and Managerial Economics in Accounting, Doctoral thesis*, (mimeo).

Hodrick, R. and. Prescott, E. (1997) Postwar U.S. Business Cycles: An Empirical Investigation *Journal of Money, Credit and Banking* 29(1) pp. 1 - 16.

Ilzetzki, E., Mendoza, E. and Végh, C. (2013) How Big (Small?) are Fiscal Multipliers? *Journal of Monetary Economics* 2(60) pp. 239 - 254.

Inderst, R. and Mueller, H. (2007) A lender-Based Theory of Collateral *Journal of Financial Economics* 84 pp. 826 – 859.

Inderst, R. and Mueller, H. (2008) Bank capital structure and credit decisions *Journal of Financial Intermediation* 17(3) pp. 295 – 314.

International Monetary Fund (2005) *Eighteenth Meeting of the IMF Committee on Balance of Payments Statistics Washington*, D.C.: IMF.

International Monetary Fund (2011) *IMF Country Report Czech Republic*, Washington DC: IMF.

International Monetary Fund (2012b) *IMF Country Report Czech Republic: Selected Issues Paper*, Washington DC: IMF.

International Monetary Fund (2015) *IMF Country Report Romania 2006 and 2014*, Washington DC: IMF.

International Monetary Fund (2012b) *IMF Country Report Montenegro*, Washington DC: IMF.

International Monetary Fund (2012c) Enhancing Financial Sector Surveillance in Low-Income Countries – Financial Deepening and Macro-Stability *Board Paper* (mimeo).

International Monetary Fund (2015) World Economic Outlook, accessed in April 2015 http://www.imf.org/external/pubs/ft/weo/2015/01/weodata/index.aspx

Jaffe, D. and Modigliani, F. (1969) A Theory and Test of Credit Rationing *The American Economic Review* 59(5) pp. 850 – 872.

Jagric, T. (2002) Business Cycles in Central and East European Countries *Journal of Eastern European Economics* 5(41) pp. 6 – 23.

Jagric, T. and R. Ovin (2004) Method of Analyzing Business Cycles in a Transition Economy: the Case of Slovenia *The Developing Economies* 1(42) pp. 42 - 62.

Jahan, S. and Mahmud, A. (2013) What Is the Output Gap? *Finance & Development* (mimeo).

Jayaraman, T. and Choong, C. (2012) Implications of Excess Liquidity in Fiji's Banking System: An Empirical Study *Munich Personal RePEc Archive Paper* No. 43505.

Jazbec, B. and Kastrati, A. (2011) Unlocking Growth Potential in the Balkans, in Nowotny, E., Mooslechner, P. and Ritzberger-Grünwald, D. (eds.): *Post-Crisis Growth and Integration in Europe*, Vienna: Edward Elgar, OeNB.

Jones, B. and Olken, B. (2005) The Anatomy of Start-Stop Growth *NBER Working Paper* No. 11528.

Kalman, R. (1960) A New Approach to Linear Filtering and Prediction Problems *Journal of Basic Engineering* 82 pp. 35 – 45.

Kalotay, K., (2001) The Contribution of Foreign Direct Investment to Transition Revisited *The Journal of World Investment* 2(2) pp. 259-276.

Kaminska, I. (2014) Why the Eurozone output gap could be a chasm Financial Times, February 2014, http://ftalphaville.ft.com/2014/02/04/1761122/why-the-eurozone-output-gap-could-be-a-chasm/

Kaminsky, G. (2005) International Capital Flows, Financial Stability and Growth *DESA Working Paper* No. 10.

Kastrati, A. (2012) Identifying Systemically Important Banks in Kosovo *Financial Stability Report No. 3*, Prishtina: CBK.

Kennedy, P. (2009) *A Guide to Econometrics, Cambridge*: Mass MIT Press.

Keynes, J. (1936) *The General Theory of Employment, Interest and Money* Cambridge: Harcourt, Brace and Company.

Khemrraj, T. (2006) Excess Liquidity, Oligopoly Banking and Monetary Policy in a Small *Open Economy New School for Social Research* (mimeo).

Khemrraj, T. (2007) Monetary Policy and Excess Liquidity: The Case of Guyana *Journal of Social and Economic Studies* 56(3) pp. 101 - 127.

Kichian, M (1999) Measuring Potential Output within a State-Space Framework *Bank of Canada Working Paper* No. 9.

Kilimani, (2009) The Link between Financial Development and Economic Growth in Uganda: A Causality Test *Paper Presented at the CSAE Conference 22–24 March* (mimeo).

Kim, J. (2004) Short Run Real Exchange Rate Dynamics: a SUR Approach *Applied Economics Letters* 11 pp. 909 - 913.

King, R. and Levine, R. (1993) Finance and Growth: Schumpeter Might be Right *The Quarterly Journal of Economics* 108(3) pp. 717 – 737.

Kleeman, L. (1995) Understanding and Applying Kalman Filtering *Department of Electrical and Computer Systems Engineering Monash University* (mimeo).

Konings, J., Rizov, M. and Vendenbussche (2003) Investment and Credit Constraints in Transition Economies: Micro Evidence from Poland, the Czech Republic, Bulgaria and Romania *Journal of Economics Letters* 78(2) pp. 253 – 258.

Koopman, S., Shepard, N. and Doornik, J.. (1998) Statistical Algorithms for Models in in State Space Form *SsfPack 3.0. OxMetrics*, London: Timberlake Consultants Ltd.

Koopman, S., Shephard, N. and Doornik, J. (2008) Statistical Algorithms for Models in State Space Form *SsfPack 3.0. OxMetrics,* London: Timberlake Consultants Ltd.

Kuttner, K. (1994) Estimating Potential Output as a Latent Variable *Journal of Business and Economic Statistics* 3(12) pp. 361 – 68.

Kydland, F. and Prescott, E. (1982) Time to Build and Aggregate Fluctuations *Econometrica* 6(5) pp. 1345 – 1370.

Lechner, M. (2011) The Estimation of Causal Effects by Difference-in-Difference Methods *Foundations and Trends(R) in Econometrics* 3(4) pp. 165 – 224.

Lee, J. (2000) The Robustness of Okun's Law: Evidence from OECD Countries *Journal of Macroeconomics* 2(22) pp. 331–356.

Levine, R. (1997) Financial Development and Economic Growth: Views and Agenda *Journal of Economic Literature* 35 pp. 688 – 726.

Long J, and Plosser C. (1983) Real Business Cycles *Journal of Political Economy* 1(91) pp. 39 – 69.

Lucas, R. (1988) On the Mechanics of Economic Development *Journal of Monetary Economics* 22 pp. 3 – 42.

Maddala, G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics* Cambridge, MA: Cambridge University Press.

Maddala, G. (1997) Introduction to Econometrics *Business and Economics* Hoboken: Wiley.

Magud, N. and Medina, L. (2011) The Chilean Output Gap *IMF Working Paper* No. 11.

Mark, N., Ogaki, M. and Sul, D. (2005) Dynamic Seemingly Unrelated Cointegrating Regressions *Review of Economic Studies* 72, pp. 797 – 820.

Milionis, A. (2003) Modelling Economic Time Series in the Presence of Variance Non-Stationarity: A Practical Approach *Bank of Greece Working Paper* No. 7.

Mishkin, F. (2001) The Transmission Mechanism and the Role of Asset Prices in Monetary Policy *National Bureau of Economic Research Working Papers* No. 8617.

Mitra, S. and Sinclair, T. (2010) Macroeconomic Fluctuations in Emerging Economies: An Unobserved Components Approach *Macroeconomic Research Group Discussion Paper Series* No 3911.

Moinescu, B, (2012) Credit Flow, Output Gap and Inflation: Nominal Convergence Challenges for the EU New Member States *Journal of Applied Quantitative Methods* 4(7) pp. 1 - 10.

Moon, H. and Perron, B. (2005) Efficient Estimation of the Seemingly Unrelated Regression Cointegration Model and Testing for Purchasing Power Parity *Econometric Reviews* 4(23) pp. 293 – 323.

Moosa, I. (1997) A Cross-Country Comparison of Okun's Coefficient *Journal of Comparative Economics* 3(24) pp. 335 – 356.

Morley, J. (2002) A State-Space Approach to Calculating the Beveridge-Nelson Decomposition *Economics Letters* 75 pp. 123 - 127.

Morley, J., Nelson, C. and Zivot, E (2002) Why are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP so Different (mimeo).

National Bank of the Republic of Macedonia (2015) Monetary and Interest Rates Statistics, accessed in April 2015 http://nbrm.mk/default-en.asp?ItemID=7869A9FDF169F7418EC8FF9457EAB055

National Bank of Poland (2015) Monetary and financial statistics, accessed in April 2015 http://www.nbp.pl/homen.aspx?c=/ascx/subgen.ascx&navid=5076

Nelson, C. and Kang, H. (1981) Spurious Periodicity in Inappropriately Detrended Time Series *Econometrica* 3(49) pp. 741 – 751.

Nelson, C. and Plosser, C. (1982) Trends and Random Walks in Macroeconomics Time Series: Some Evidence and Implications *Journal of Monetary Economics* 10 pp. 139 - 162.

Nelson, R. (1989) On the measurement of capacity utilisation *Journal of Industrial Economics* 3(37) pp. 273 - 286.

Neumeyer, P. and Perri, F. (2004) Business Cycles in Emerging Economies: The Role of Interest Rates *NBER Working Paper* No. 10387.

O'Connell, S. (2005) Note on Excess Liquidity, IMF (mimeo).

OECD (2010) Counter-Cyclical Economic Policy OECD Economics Department Policy Notes No. 1.

OECD (2015) Business Cycle Analysis, accessed in July 2015 http://stats.oecd.org/mei/default.asp?rev=2

Ogunc, F. and D. Ece (2004) Estimating the Output gap for Turkey: an Unobserved Components Approach *Applied Economics Letters* 3(11) pp. 177 – 182.

Okun, A. (1962) Potential GNP: Its Measurement and Significance *American Statistical Association Proceedings of the Business and Economic Section,* Washington, D.C.: American Statistical Association pp. 98 – 103.

Ooms, M. (2007) Panel Data, SUR and GLS *Tinbergen Institute Amsterdam* (mimeo). http://www.feweb.vu.nl/econometriclinks/orientatie/EconometricsII20079.pdf

Orts, V., Alguacil, M. and Cuadros A., (2008) The Last Enlargement of the EU and the Effects of Inward Foreign Direct Investment: A Dynamic Panel Data Approach *Review of Development Economics* 12(3) pp. 594 - 604.

Parks, R. (1967) Efficient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated *Journal of the American Statistical Association* 62 pp. 500 – 509.

Peia, O. and Roszbach, K. (2011) Finance and Growth: The Good, the Bad and the Causality (mimeo).

Perron, P. (1986) Tests of Joint Hypotheses for Time Series Regression with a Unit Root *University of Montreal* (mimeo).

Perron, P. (1989) The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis *Econometrica* 6(57) pp. 1361 – 1402.

Peter K. (1987) The Cyclical Component of U.S. Economic Activity *Quarterly Journal of Economics* 102 pp. 797-814.

Petkovska, B. and Kabashi, K. (2011) Estimation of Potential Output and Output Gap in Macedonia According to Different Estimation Methods *National Bank of the Republic of Macedonia Working Paper*.

Perloff, J. and Wachter, M. (1979) A Production Function - Nonaccelerating Inflation Approach to Potential Output: Is Measured Potential Output Too High? *Carnegie - Rochester Conference Series on Public Policy* 10 pp.113 - 163.

Perman, R and Tavera, C. (2005) A Cross-Country Analysis of the Okun's Law Coefficient Convergence in Europe *Applied Economics* 21(37) pp. 2501 – 2513.

Phillips, A. (1958) The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957 *Economica* 100(25) pp. 283 – 299.

Phillips, P. and Perron, P. (1988) Testing for a Unit Root in Time Series Regression *Biometrika Trust* 2(75) pp. 335-346.

Pistor, K., Raiser, M. and Gelfer, S. (2000) Law and Finance in Transition Economies *EBRD Working Paper* No.48.

Plosser, C. (1989) Understanding Real Business Bycles *Journal of Economic Perspectives* 3(3) pp. 51-77.

Podpiera, J. and Weill, L. (2008) Bad Luck or Bad Management? Emerging Banking Market Experience *Journal of Financial Stability* 4(2) pp. 135 – 148.

Polleit, T. and Gerdesmeier, D. (2005) Measures of Excess Liquidity *Frankfurt School of Finance and Management Working Paper Series* No. 65.

Pomenkova, J. and Marsalek, M. (2011) Time and Frequency Domain in the Business Cycle Structure *MENDELEU Working Papers* No. 7.

Primus, K. (2013) Excess Reserves, Monetary Policy and Financial Volatility *Munich Personal RePEc Archive* No. 51670.

Pruskyi, J. and Szpunar, P. (2009) Capital Flows and Their Implications for Monetary and Financial Stability: the Experience of Poland *Bank for International Settlements Papers* No. 44.

Pyo, H. (2008) The Estimation of Industry-level Capital Stock for Emerging-Market and Transition Economies *The 2008 World Congress on National Accounts and Economic Performance Measures for Nations,* Washington DC.

Quantitative Micro Software (2010) *EViews 7 User's Guide II,* Irvine CA: EViews.

R.A. (2013) What Does "Above Potential" Mean? The Economist August 2013, http://www.economist.com/blogs/freeexchange/2013/08/business-cycles

Rachdi, H. and Mbarek, H. (2011) The Causality Between Financial Development and Economic Growth: Panel Data Cointegration and GMM System App roaches International *Journal of Economics and Finance* 1(3) pp. 143 - 151.

Raiffeisen Bank International (2015) *Annual Reports* 2005, 2008, 2011 and 2013, Vienna: RBI.

Rand, J. and Tarp, F. (2001) Business Cycles in Developing Countries: Are They Different? *CREDIT Research Paper* No. 21.

Reininger, T., Schardax, F., and Summer, M. (2002) Financial System Transition in Central Europe: The First Decade *Société Universitaire Européenne de Recherches Financières.*

Roaf, J., Atoyan, R., Joshi, B., and Krogulski, K., (2014) 25 Years of Transition Post-Communist Europe and the IMF, Regional Economic Issues, *International Monetary Fund IMF Special Report.*

Robinson, J. (1963) Essays in the Theory of Economic Growth *The Economic Journal* 290(73) pp. 295 – 297.

Room, M. (2001) Potential Output Estimates for Central and east European Countries Using Production Function Method *Working Papers of Eesti Pank* No. 2.

Ross, L., Loyaza , N. and Beck, T. (1999) Financial Intermediation and Growth: Causality and Causes *World Bank Policy Research Working Paper* No. 2059.

Roubini, N. (2011) Eurozone Crisis: Here Are the Options, Now Choose *EconoMonitor,* November, 2015. http://www.economonitor.com/nouriel/2011/11/09/eurozone-crisis-well-at-least-we-have-options/

Roubini, N. (2015) Unconventional Truth *Project Syndicate,* February, 2015. http://www.project-syndicate.org/commentary/unconventional-monetary-policies-and-fiscal-stimulus-by-nouriel-roubini-2015-02

Ruffer, R. and Stracca, L. (2006) What is Global Excess Liquidity and Does It Matter ? *ECB Working Paper* No. 696.

Sacerdoti, E. (2005) Access to Bank Credit in Sub-Saharan Africa: Key issues and Reform Strategies *International Monetary Fund Working Paper* No. 166.

Shepherd, D. and Driver, C. (2003) Inflation and Capacity Constraints in Australian Manufacturing Industry *The Economic Record* 79(245) pp. 182 – 195.

Saxegaard, M. (2006) Excess Liquidity and Effectiveness of Monetary Policy: Evidence from Sub-Saharan Africa *IMF Working Paper* No. 115.

Solow, R. (1956) A Contribution to the Theory of Economic Growth *Quarterly Journal of Economics 70(1)* pp. 65 – 94.

Sorensen, C. and Werner, T. (2006) Bank interest rate Pass-Through in the Euro Area A Cross-Country Comparison *European Central Bank Working Paper* No. 580.

Sramkova, L., Kobilicova, M. and Krajcir, Z. (2010) Output Gap and NAIRU Estimates within State-Space Framework: An Application to Slovakia *Financial Policy Institute, Economic Analysis* No. 16.

StataCorp (2013) *Stata 13 Base Reference Manual*, Texas: Stata Press.

State Bank of Latvia (2015) Monetary Statistics, accessed in April 2015 https://www.bank.lv/en/statistics/monetary-statistics/key-indicators-of-the-mfi-banking-sector

State Space using SsfPack 2.2 (1999) *Econometrics Journal* 1(2) pp. 107 – 160.

Statistical Agency of Kosovo (2015) National Account, Labour Market and Prices, accessed in October 2011 and February 2014 http://ask.rks-gov.net/eng/

Stiglitz, J. and Weiss, A. (1981) Credit Rationing in Markets with Imperfect Information *Journal of American Review* 3(71) pp. 393-410.

Svejnar, J. (2002) Transition Economies: Performance and Challenges *Journal of Economic Perspectives* 16(1) pp. 3-28.

Stock, J. and Watson, M. (2015) Core Inflation and Trend Inflation (mimeo).

Studenmund, A. (2006) Using Econometrics: a Practical Guide *Addison-Wesley Series in Economics*, Boston: Addison Wesley Pearson.

Svensson, L. (1997) Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets *European Economic Review* 41(6) pp. 1111 – 1146.

Szorfi, B. (2011) Estimating Underlying Wage Dynamics Using Multivariate Kalman Filter *National Bank of Hungary* (mimeo).

Takats, E. and Upper, U. (2013) Credit and Growth after Financial Crises *Bank for International Settlements Working Paper* No. 416.

Taylor, J. (1992) Price Stabilization in the 1990s: An Overview *Monetary and Economic Studies*, *Institute for Monetary and Economic Studies*, *Bank of Japan*, vol. 1(10) pp. 35-43.

The Central Bank of Hungary (2015) Statistical Time Series, accessed in April 2015 http://english.mnb.hu/Statisztika/data-and-information/mnben_statisztikai_idosorok

The Central Bank of Republic of Kosovo (2015) Statistical Time Series, accessed in April 2015 http://bqk-kos.org/?m=t&id=55

Thompson, S., Sul, D. and Bohl, M. (2002) Spatial Market Efficiency and Policy Regime Change: Seemingly Unrelated Error Correction Model Estimation *American Journal of Agricultural Economics* 84 pp. 1042 - 1053.

Tsalinski, T. (2007) Two Approaches to Estimating the Potential Output of Bulgaria *Bulgarian National Bank Discussion Paper* No. 57.

Verbeek, M. (2000) *Modern Econometrics*, New York: Wiley.

Vogelvang, B. (2005) *Econometrics: Theory and Applications with EViews*, London: Financial Times Prentice Hall.

Vroey, M. and Malgrange, P. (2011) The History of Macroeconomics from Keynes's General Theory to the Present *Institut de Recherches Economiques et Sociales de l'Universite Catholique de Louvain Discussion Papers* No. 28.

Waller, C. and Lewarne, S. (1994) An Expository Model of Credit Rationing *Journal of Macroeconomics* 3(16) pp. 539 – 545.

Watson, M. (1986) Univariate Detrending Methods with Stochastic Trends *Journal of Monetary Economics* 1(18) pp. 49 – 75.

Welch, G. and Bishop, G. (2001) An Introduction to the Kalman Filter *University of North Carolina at Chapel Hill* (mimeo).

Winship, C. (2009) Lecture 16: Generalized Least Squares *Sociology 203a-09* (mimeo).

Wooldridge, J. (2002) *Econometric Analysis of Cross Section and Panel Data*, Cambridge: Mass MIT Press.

Wooldridge, J. (2010) *Econometric Analysis of Cross Section and Panel Data*, 2nd edition, Cambridge: Mass MIT Press.

World Bank (2015) *Doing Business Report 2006 - 2014*, Washington DC: The World Bank.

World Bank (2015) Financial Sector Indicators, accessed in April 2015 http://data.worldbank.org/topic/financial-sector

World Bank (2015) Doing Business Indicators 2006 - 2014, accessed in April 2015 http://www.doingbusiness.org/rankings

Zaman, C.(2002) The Output Gap and the Potential GDP of the Yugoslav Economy *Policy and Legal Advice Centre-An EU-funded project managed by European Agency for Reconstruction* (mimeo).

Zhang, C. and Murasawa, Y. (2011) Output Gap Measurement and the New Keynesian Phillips Curve for China Economic Modelling *Elsevier* 6(28) pp. 2462 - 2468.

Zellner, A. (1979) Causality and Econometrics, in: K. Brunner and A.H. Meltzer (eds.), *Three Aspects of Policy and Policymaking* Amsterdam: Elsevier, North-Holland.

Zivot, E. and Andrews, D. (1992) Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis *Journal of Business and Economic Statistics* 3(10) pp. 251 - 270.