VILNIUS UNIVERSITY

AGNĖ REKLAITĖ

MEASURING THE GLOBALISATION EFFECT: THE APPLICATION OF A DYNAMIC HIERARCHICAL FACTOR MODEL

Doctoral dissertation

Physical sciences, mathematics (01 P)

Vilnius 2016

This dissertation was carried out at Vilnius University in the period of 2011–2015.

Scientific supervisor:

prof. dr. Marijus Radavičius (Vilnius university, physical sciences, mathematics — 01 P)

VILNIAUS UNIVERSITETAS

AGNĖ REKLAITĖ

HIERARCHINIO DINAMINIO FAKTORINIO MODELIO TAIKYMAS GLOBALIZACIJOS EFEKTUI VERTINTI

Daktaro disertacija

Fiziniai mokslai, matematika (01 P)

Vilnius 2016

Disertacija rengta 2011–2015 metais Vilniaus universitete.

Mokslinis vadovas:

prof. dr. Marijus Radavičius (Vilniaus universitetas, fiziniai mokslai, matematika — 01 P)

Ackowledgements

First and foremost I would like to thank my advisor professor Dr. Marijus Radavičius for his guidance in my research, immense knowledge and persistent encouragement. I would also like to express gratitude to my colleagues Dr. Virmantas Kvedaras and Dr. Dmitrij Celov for valuable discussions and suggestions. I am grateful to reviewers Dr. Kęstutis Dučinskas and Dr. Algirdas Bartkus for insightful comments and advice. I greatly appreciate the help of Prof. Alfredas Račkauskas dealing with difficulties in my studies.

My sincere thanks goes to my high school mathematics teacher Dr. Antanas Skūpas for enthusing me with mathematics. I am also grateful to Dr. Vaidotas Zemlys-Balevičius for introducing me to factor modelling.

I am grateful to my family for their support. I am indebted to my mother and my in-laws for helping me find time for research. Special thanks goes to my daughters Vaiva and Milda for being sweet and patient children, and making it possible for me finish what I started. Finally, I would like to express my appreciation to my husband Jonas, who encouraged me to embark upon this academic adventure and supported me in every stage of graduate studies.

Contents

Abbreviations

2FM – Dynamic 2-factor model

CEI – Coincident economic index

CSGR – Centre for the Study of Globalisation and Regionalisation

DLM – Dynamic linear model

DHFM – Dynamic hierarchical factor model

EC – European Community (former European Economic Community)

EFTA – European Free Trade Association

FDI – Foreign direct investment

GDP – Gross Domestic Product

GNP – Gross National Product

ISM – Institute of Supply Management

KOF – Konjunkturforschungsstelle (Economic Research Center in Zurich)

LDC – Less developed country

MAE – Mean absolute error

MAPE – Mean absolute percentage error

MCMC – Markov chain Monte Carlo

ME – Mean error

MPE – Mean percentage error

NBER – National Bureau of Economic Research (in U.S.)

NGI – New globalisation index

OECD – Organisation for Economic Cooperation and Development

OLS – Ordinary least squares

PCA – Principal component analysis

RMSE – Root mean squared error

TFP – Total factor productivity

TRI – Trade restrictiveness index

UN – United Nations

Notations

$\mathbf{0}$ – vector or matrix of zeroes

$\boldsymbol{I_m}$ – m-dimensional identity matrix

$\propto$ – proportional to

Δ – difference operator

$\mathbf{E}$ – expectation

L — lag operator

$\mathbb{R}^+$ — positive real number set

$\mathbb{R}^n$ — n-dimensional space of real numbers

$\mathbb{1}_{\{x \in A\}}$ — indicator function denoting if x belongs to the set A

RPL_j — relative price level for country j

P_j — consumption price index for country j

r — exchange rate

$\mathcal{N}(m, \sigma^2)$ — normal distribution with mean m and variance σ^2

C_t — coincident economic index (latent)

λ — a vector of loadings

F_t — single factor denoting the unobserved state of economy

$G_{b,t}$ — block level factors: domestic $G_{1,t}$ and foreign $G_{2,t}$

X_t — a vector of coindicent indicators

μ_t — serially correlated error term

$D(L)$ — matrix lag polynomial describing autoregressive structure of μ_t

ψ — autoregressive coefficient for F_t

$\epsilon_t, \varepsilon_t, u_t$ — serially independent error terms

$d(x, y)$ — distance function between x and y (metric)

Λ_G — block-level loadings on block-level factors G_{bt}

Λ_F — loadings of block-level factors on common factor F_t

X_{bit} — series of leading indicators structured into blocks $b = 1, 2$

$Y_t = \frac{C_{t+2}}{C_t}$ — growth of coincident economic index (scaled to have 0 mean and unit variance)

$\alpha_{1,t}$ — coefficient denoting the amount of growth explained by domestic factor on the future growth of economy

$\alpha_{2,t}$ — coefficient denoting the amount of growth explained by foreign factor on the future growth of economy

c — mean of $\alpha_{1,t}$ process

γ — the restriction parameter for the sum $\alpha_{1,t} + \alpha_{2,t}$

ϕ — autoregressive coefficient for $\alpha_{1,t}$

$F_{1,t}$, $F_{2,t}$ — factors from non-structural approach, extracted from leading indicators

Introduction

Technological progress and the increasing accessibility to information enables geographically distant factors to have an effect on economic and socio-cultural processes. Similar patterns are observed in econometric modelling: the traditional domestically oriented macro-econometric models struggle with diminishing accuracy and globalisation is increasingly addressed as the underlying cause of this phenomenon. An application of the extended Conference Board methods by Drechsel and Sheufele [1] shows that more and more indicators have to be included in construction of the leading economic index to keep up with the accuracy of previously constructed models. This result could indicate that economic processes are becoming of a more complicated structure impelled by the increasing amount of information available for a single agent of economy and therefore affecting its decision-making [2]. The accuracy of domestically oriented models deteriorates with time and this issue is addressed by Fichtner et al. [3]. Their findings suggest that it is caused by globalisation, thus adding information about the external environment improves the forecast performance.

Selecting the best indicators for domestically oriented macro-econometric models is demanding, and expanding the potential indicator list to include foreign variables brings a new challenge: finding a way to include the most relevant information and to sustain statistical feasibility of the model. There are two ways to address this problem: either find a small number of international indicators to include into the model directly, or use the approach that is suitable for forecasting using a large number of predictors, such as factor modelling [4–7]. Eickmeier and Ziegler [8] analysed models of output prediction (52 studies) and concluded that data-rich methods outperform small-scale models.

Factor models have been proven to be consistent and asymptotically efficient [9] and they are frequently used for short-term macroeconomic forecasting [10–14]. However, they are criticised for the lack of interpretation since the factors are extracted from large data sets without taking into account the structure of the data. In response to that issue, dynamic hierarchical factor models were offered by Moench, Ng and Potter [15]. This method uses a structural approach which provides the basis for interpretation, but it has not found much practical use yet, and has been applied mostly for inspecting how much of the total variance could be explained by the structure of the data [16].

Ever since Keynes referred to econometrics as *statistical alchemy* [17], econometricians have been making methodological contributions to convince the sceptics of scientific nature of econometrics [18], in spite of lack of possibility to have controlled economic experiments. Rapid progress of this process is dependent on the ability to employ the scientific method. The advancement of structural econometrics expands the prospects to validate economic hypotheses, thus dynamic hierarchical factor models could be employed for methodological augmentation through enabling the validation of more abstract hypothesis than in the case of regression.

A good globalisation measure would contribute to general discussions on globalisation and its effect in various areas of interest. It would also be beneficial for econometricians to plan the updates and revisions of marco-econometric models. This doctoral dissertation addresses the problem of how to measure a latent process such as globalisation which manifests itself in a large number of statistical indicators. The accessibility to information conditions the advancement of globalisation, however it is spurred only if the knowledge is implemented in practice, therefore the globalisation in this thesis is measured through its impact on the focal economy.

Topicality of the research

Globalisation presents in various forms. The on-line retailing is gaining more popularity: digital buyer penetration has reached 24.3% of the global population in 2015 and is forecast to reach 32.8% by the year 2019 [19]. The worldwide number of international parcels has doubled in the period of 2007–2014 [20]. Also, the widespread opening of new commercial venues increases the global presence of large companies, such as McDonald's, IKEA, H&M etc., and it contributes to the processes of international integration.

The manifestation of the globalisation effect also presents in macroeconometric models. This phenomenon requires to update the models and to include the supra-national element. The inclusion of international indicators into macro-econometric models gives a boost in accuracy [1, 3, 4]. However, the pace of globalisation is likely to increase: with the progress of information technology the communication with distant parts of the world is becoming more available and prompt. Nevertheless, even if the foreign component is included into a macro-econometric model it might become outdated because the globalisation is gaining momentum.

Considering this issue it is important to assess the globalisation effect and to evaluate the development of this phenomenon. For that reason a measure of globalisation is needed which has a quality of comparability across different time points, stems from economic domain and has an ability to capture developments in different sectors.

Objective and tasks

The main goal of this study is to propose a new globalisation measure and to develop a method to measure a latent phenomenon whose effect can be gauged in a large number of statistical indicators. The proposed measure should have the clarity in what it measures and the ability to capture multifaceted nature of the phenomenon, thus having superiority over previously constructed measures. The notion that globalisation and openness spur economic growth is used as a core assumption and the idea of measurement is to quantify the portion of growth explained by international variables relative to domestic ones. In order to achieve that the following tasks were formed:

1. Construct an indicator of economic growth which reflects the multi-domain developments across different sectors of economy.

2. Build the dynamic hierarchical factor model, describing latent leading domestic and foreign factors and their relationship to each other as well as their linkage to directly measured indicators.

3. Assess the time-varying domestic and foreign loadings on future growth of economy and use the results to derive the globalisation index.

In the light of discussions on the globalisation and its pace of development, a hypothesis was formed in order to assess the practical use of this new measure: the portion of economic growth explained by foreign indicators should be increasing over time. This hypothesis was validated applying the new measure on Lithuanian data.

Research methods

Factor models are the basis of applied reseach methods. Dynamic hierarchical factor model and dynamic linear model were used to produce main results. Models were evaluated using Bayesian econometric approach — with Gibbs sampling algorithm on Markov chain Monte Carlo. The simulations were carried out employing statistical software *R* and its package *dlm*. The parameters of a dynamic linear model were evaluated using maximum likelihood method. The selection of leading indicators employed hierarchical clustering methods and the least angle regression algorithm.

Scientific novelty

A new measure of globalisation was designed which has the quality of reflecting the multi-domain developments of international integration and was built using a statistically sound technique by introducing the factors acquired from a dynamic hierarchical factor model into a dynamic linear model. Upon developing the new measure of globalisation, the approach of using *a priori* knowledge was taken in order to design the structure of the model fitted to the structure of the data. This approach gives a set-up for the interpretation of results. Also this expands the possibilities to employ the scientific method in econometrics: formulating the hypothesis before the empirical analysis and validating it after; the design of the model enables validation methods for more abstract hypotheses than using a regression.

Statements presented for defence

1. A new globalisation measure is proposed which is based on measurement of the load of foreign variables on the forecast economic growth.
2. The methodology for evaluating the presented globalisation measure is developed, which enables the evaluation of impact of grouped indicators on the variable of interest.
3. Presented method fits the design of a dynamic hierarchical factor model to suit the structure of the data and produce a set-up for interpretation of the results.
4. The portion of future growth explained by foreign indicators relative to domestic ones is increasing for the Lithuanian economy and it reflects the globalisation effect.

Publications by the author

Articles

1. A. Reklaitė. "Coincident, Leading and Recession Indexes for the Lithuanian Economy", *Baltic Journal of Economics*, *11(1)*, 2011, p. 91–107.

2. A. Reklaitė. "Globalisation effect measure via hierarchical dynamic factor modelling", *Equilibrium*, *10(3)*, 2015, p. 139–149.

3. A. Reklaitė. "Measuring foreign impact: leading index construction using hierarchical dynamic factor model", *Baltic Journal of Economics*, *16(1)*, 2016, p. 21–32.

Presentations in conferences

1. A. Reklaitė. *Coincident and leading indexes construction for a small open economy*, XI$^{\text{th}}$ International Vilnius Conference on Probability Theory and Mathematical Statistics. 30 June - 4 July, 2014, Vilnius, Lithuania

2. A. Reklaitė. *Globalisation effect measure via hierarchical dynamic factor modelling*, VIII$^{\text{th}}$ International Conference on Applied Economics. Contemporary Issues in Economy. June 18-19, 2015, Toruń, Poland

1. Review of the Literature

1.1. Globalisation

1.1.1. Introduction

Globalisation is a concept that has been addressed to by the media, politicians and various areas of academia, especially the social sciences. This process affects many different sociocultural issues such as usage of languages [21–23], tourism [24–26], organised crime [27–29], business management [30, 31] and many others. Economics covers a lot of globalisation related research topics and addresses its links to various economic phenomena, e.g. the changes in inequality [32–34], inflation [35–37], economic growth [38, 39].

Even though the term is widely recognised, many authors have struggled to define it. Sometimes it is defined as the increasingly close international integration of markets for goods, services and factors of production, labour and capital [40]. Rennen and Martens [41] describe it as an interactive co-evolution of millions of technological, cultural, economic, social and environmental trends at all conceivable spatio-temporal scales. Albrow [42] notes that globalisation refers to all those processes by which the peoples of the world are incorporated into a single world society, a global society. What unifies the different definitions is a reference to the globalisation as a set of multi-domain developments of international integration, therefore the effects of the globalisation should manifest in numerous different indicators.

One aspect of globalisation is the convergence in prices of traded commodities and services [40], this process is largely affected by the international competition through development of IT sector and the increase in on-line trading. Thus the globalisation is spurred by the technological advance and its penetration into economic activities. With the shift towards knowledge economy this process is expected to gain more momentum and bring the globalisation to new heights.

Several theories note the importance of technological spillovers and the international transmission of knowledge as a source of growth for open economies. Developed countries' economic growth relies on advance of a knowledge based services sector and the innovation [43]. Since investments in technology drive the growth by creating broader productivity gains in the form of economy wide spillovers [44] the growth of developed countries should affect the other economies. A study by Schneider [45] revealed that the foreign technology has a stronger impact on per capita GDP growth than the domestic technology. Similar results were achieved by Lee [46]: he found that an open economy has a higher growth rate of income if foreign capital goods are used relatively more than domestic capital goods for the production of capital stock. The findings by Harrison [47] show that there is a significant positive relation between openness of economics and economic growth. These results indicate that foreign variables affect the growth of economics and with the technological advance stipulating the globalisation, the foreign effect on economic growth is conditioned to increase. This effect should be particularly conspicuous on small open economies. In order to monitor and ascertain it a measure of globalisation is needed.

1.1.2. Measures of globalisation

Globalisation is a complex phenomenon and it manifests across many different sectors. Any measures of globalisation are dependant on the definition and information or data available and therefore are embedded in certain dimensions. Evaluation of globalisation effect relies heavily on quantitative variables, therefore not every aspect of this phenomenon can be captured by measurement.

Trade-based measures

International trade is an obvious feature of globalisation. It is natural to assume that economies with more intense international trade are more affected by globalisation. Therefore the openness of the economy corresponds to the measure of globalisation in the dimension of trade. A group of researchers embraced this idea and constructed their estimates. It is noteworthy that the term "openness" and the measure of it is usually analysed in the context of trade policy. However, it is also informative about international integration in the dimension of trade.

The idea that free trade should encourage economic growth was first brought up by Adam Smith [48]. He argued that free trade should cause the commodity prices to converge across different countries. As a result, each country would specialise in producing the merchandise where it has a superiority and the productivity would increase because of economies of scale. This basic idea is covered by voluminous research by many economists and the relationship between the openness of trade policy and either economic growth or degree of specialisation was examined in many different ways.

An example of latter approach is a study by Quah and Rauch [49]. Using a model of endogenous growth and trade shares as an indicator of openness policy they showed that an increased openness to trade can lead to an increased specialisation through learning by doing. Whether specialisation accelerates productivity growth was later explored by Weinhold and Rauch [50] in an empirical study where specialisation was measured by Herfindahl [51] index for the manufacturing sector. They used a panel linear model with fixed effects and found a positive relationship between the specialisation and the manufacturing productivity growth.

More popular way to test the relation between open trade and economic growth was to inspect the economic growth as measured by gross domestic product (GDP) or gross national product (GNP) and some sort of openness measure.

Early attempts to investigate the link between the openness of trade policy and the economic growth produced a handful of different measures derived from variables of trade. In 1977 Michaely [52] used the rate of the change of the ratio of exports in the total product as a measure of openness. In 1978 Balassa [53] used 3 measures of trade policy, derived from international trade: export growth rates, the absolute increment in exports to the absolute increment in GNP and the ratio of exports to GNP. His findings show that those measures are sensitive to specification, e.g. the selection of the base year or a version of GNP (if it is per capita or not), but results are very similar for majority of countries. In his 1985 paper Balassa [54] introduced a correction for the openness indicator: a trade orientation variable that adjusts for natural resource availabilities. This adjustment significantly improved the results for countries with the large extent of exports of oil products in the 1970s.

Helliwell and Chung [55] used a five-year moving average ratio of foreign trade to GDP as an indicator of openness to inspect its effect on the pace of international convergence of labour efficiency as measured by real output per worker. The functional form of their hypothesis implies that it is the proportionate changes in the trade share that affect the productivity level, and that the equilibrium efficiency level will be unaffected by the level of the equilibrium trade share. Their findings validate the hypothesis and show the strongest effect in European countries.

In 1995 Warner and Sachs [56] defined their openness indicator as a binary variable, which was set to zero for a specific country if any of five conditions, describing various trade limitation features, were met:

1. The country had average tariff rates higher that 40%

2. Country's non-tariff barriers corresponded on average to more than

40% of imports

3. Country a had socialist economic system

4. Country had a state monopoly on major exports

5. Country's black-market premium exceeded 20%

This indicator has some strong features, such as robustness to the specification, however it is not dense in information since it can take only 2 values and some of the criteria are getting outdated (e.g. (3)). Even though those 5 indicators are economically justified, it has been observed by Rodriguez and Rodrik [57] that variables (2), (3) and (4) are uninformative since index produced using only (1) and (5) conditions is almost identical to Warner Sachs index. Another drawback of this measure is that for many countries some indicators are available only at 1 time point, therefore inspection of evolution of openness is impossible.

Leamer [58] offered to measure the openness using a difference in predicted and observed trade intensity ratios. He used an empirical Heckscher-Ohlin [59] model to estimate net trade flows and trade intensity ratios for 183 commodities for 53 countries. His method later was expanded by Wolf [60] using larger set of factors of production and more disaggregated categories for commodities. The constructed indexes measured the distance between the actual trade and the predicted trade by the model under the conditions of free trade.

The idea that the trade openness can be derived from difference between actual trade values and predicted values by a model was also adopted by Lee [61]. He regressed import shares on the land area, a distance from major trading partners, import tariffs, and black-market premia, and then calculating the predicted value of imports when the actual values of tariffs black-market premia are replaced by zeros.

Frankel and Romer [62] noted that the international trade is influenced by geography to a large extent and measures that ignore geographical parameters of countries tend to overvalue the openness for countries that are close to populous economies (e.g. Belgium) and undervalue it for countries that are geographically far from them (e.g. New Zealand). Thus, they developed instrumental variables from geographical parameters that correspond to the portion of the trade that is accounted for by purely geographical conditions.

In 2009 Naghshpour and Sergi [63] used the share of international trade, i.e. the sum of imports and exports, to GDP and scaled the acquired ratios to produce Z-scores. Next, the countries were classified into 6 categories according to their Z-scores, selecting the endpoints of intervals at integer numbers. This simplistic approach is vulnerable to the same critique that was aimed at trade-based openness measures, and the authors' claim that it is the first attempt to construct a meaningful and statistically sound globalisation index seems overstated. However, even though authors failed to mention it, one thing that can be noticed from their results – the distributions of calculated Z-scores evolve over time so that kurtosis is diminishing and the values are more and more concentrated near zero. This is evidently a result of globalisation and it implies that countries are becoming more similar in the dimension of trade.

Price-based measures

Findlay and O'Rourke [64] note that changes in trade do not necessarily have a connection to globalisation: it could be caused by changes in supply and demand. Therefore a convergence in commodity prices would be a more accurate measure of globalisation.

In an influential paper by Barro [65] relative domestic price of investment goods to international prices were used as an indicator for market distortions. These indicators were included in a neoclassical growth model and the results revealed an inverse relationship between country's per capita growth and initial income levels once human capital is controlled by education indicators and fertility rates. The findings were also suggestive of distortions for investment goods to be adverse for growth.

The comparative cross-country prices were used to measure the outward-orientation by Dollar [66]. He adopted the expression of country's index of relative price (to U.S.) level $RPL_j = 100 \times \frac{r \cdot P_j}{P_{U.S.}}$ by Summers and Heston [67] (where r is exchange rate and P_i is the consumption price index for country j) and regressed it on country's endowments. The acquired residuals were averaged over 10 year period and the acquired index indicates the magnitude to which a country's prices are high or low, given its endowments. Dollar concluded that a country sustaining a high price level over many years would clearly have to be a country with a relatively large amount of protection. Even though Dollar's measure of outward-orientation gained a huge popularity among economists analysing international trade, this method has been criticised for adopting unlikely assumptions. Rodriguez and Rodrik [57] pointed out that this measure works well only if all trade barriers are on the import. However the restrictions on export are also applied in practise by many countries. Another weakness of this method might be a lack in accuracy since this measure may be influenced by transportation costs, monetary and exchange rate policies.

Trade barrier based measures

Another alternative to measure openness is based on trade restrictiveness and is constructed using data on tariffs and quantitative trade restrictions. Anderson and Neary [68] proposed a trade restrictiveness index (TRI). TRI is welfare-equivalent and it is derived from comparison of two equilibrium conditions: one with free trade and another with imposed tariffs and trade quotas. They derive a uniform tariff equivalent in terms of the welfare to a set of trade restrictions from data set. The biggest advantage of this approach is that it is firmly based on economic theory and the resulting index is derived from assumptions and data using equilibrium conditions. Unfortunately, this method requires data to have a certain variety of indicators and therefore this index could not be evaluated for a large portion of countries.

Proxy variables to openness

In studies examining the connection between the openness and the economic growth the ratio of population to total area have been used as a proxy for openness under the assumption that counties with high population density tend to be more open than others [56]. Another proxy variable for trade openness – black market premium for foreign exchange was used by Levine and Renelt [69], Barro and Lee [70], Harrison [47]. The theoretical justification for it is that foreign exchange restrictions act as a trade barrier under certain conditions.

Memberships in trade organisations, such as EC, EFTA, have also been used as indicators for global or regional integration [71].

Composite indicators

Many of the mentioned openness measures have their own strengths and weaknesses which have been debated over years, and attempts have been made to improve them. Edwards [72] brought up an argument that constructing superior measure of openness is not as important as comparison of the results, using the existing measures. He used 9 openness indexes:

1. Warner and Sachs openness index [56]
2. World Development Report Outward Orientation Index, developed by Dollar [66]
3. Leamer's Openness Index [58]

4. The average black-market premium (this measure has been used as an openness indicator by Levine and Renelt [69])

5. Average import tariff on manufacturing

6. Average coverage of non-tariff barriers

7. The Heritage Foundation index on distortions in international trade [73]

8. Collected trade taxes ratio. This index was calculated by author as the average ratio of total revenues on taxes on international trade to total trade

9. Wolf's index [60] of imports distortions

The findings show that these 9 indicators tell the same story. Edwards concluded that first principal component calculated from indicators 1, 4, 5, 6, 9 is the most informative measure (it explains 60% of variation) and he used it to examine the openness relation to the growth of economy.

Wacziarg [74] constructed a composite indicator of trade policy openness from 3 indicators: the average import duty rate, the non-tariff barriers coverage ratio, and the Warner-Sachs binary indicator of openness. Weights used to construct the combined index were acquired from a regression of trade ratio to GDP on these three indicators plus some gravity indicators, such as log of land area and log of population, as well as the growth rate of per capita GDP.

In 2001 World Market Research Centre presented G-index, developed by Randolph [75], which meant to measure globalisation defining it as "the ever closer knitting together of a one-world economy". Therefore the 90% of indicators corresponded to economical integration and the rest were attributed to technology: 5% for telephone traffic and the 5% weight for internet hosts. The calculation method is weighted summing, weights were based on author's insight with 70% load on international trade and service exports. This index is appealing for its wide coverage – 185 countries, some of them having time series of 30 points length. It was pointed out by Martens and Zywietz [76], that this index favours small trading nations that have huge (transit) trade volumes with respect to their internal economy.

The first attempt to construct a quantitative globalisation index which was not an openness measure and did capture activities in different domains was by Kearney [77] in 2001. The indicators used covered areas of trade, foreign direct investment, portfolio capital flows, income payments and receipts, international travel and tourism, international telephone traffic, cross border transfers, number of internet users, internet hosts and secure servers, number of international organisations, UN Security Council missions participated in, number of foreign embassies. The method for combining them was weighted summing and the most tedious task before applying it, is adjusting the indicators so that they are comparable across countries and normalising to get them into the same measurement scale. The weights were chosen according to the author's beliefs therefore this index is somewhat subjective. Kearney index in cooperation with Foreign Policy Magazine was updated annually until 2006 and the weights were revised, therefore it is possible to find several Kearney index estimates for the same period that differ dramatically for some countries from different editions. This index is also criticised for not being clear of what exactly it measures and that indicators from different countries are calculated using different methodologies therefore not possessing the desired feature of cross-comparability [78].

The Centre for the Study of Globalisation and Regionalisation (CSGR) globalisation index was developed in 2005 by Lockwood and Redoano [79]. It was designed as complementary to Kearney index. The indicators used to construct it are of the same variables that are used in Kearney index. The main improvement over the Kearney index is statistically sound weighing procedure. The authors evaluate the weights using principal component analysis (PCA), the weights extracted from the first principal component.

KOF index of globalisation proposed by Dreher [80] uses data that largely intersect with indicators used by Kearney [77]. The additional indicators were on trade restrictions, also the data pool of information flows and social connectivity was expanded by adding extra indicators, such as foreign population, cable television and number of McDonald's restaurants. The indicators were organized in a hierarchical fashion to produce one globalisation index and 3 sub-indexes on economic integration, social globalisation and political engagement. The weights to combine the indicators were acquired using principal component analysis and selecting the loadings from the first principal component. KOF index is U.S.-centric on some level as several indicators are clearly favoured in U.S. (number of McDonald's restaurants or telephone average costs of call to USA), which might be considered subjective since the design of this index pre-sets the U.S. and Canada to be on the top of the list.

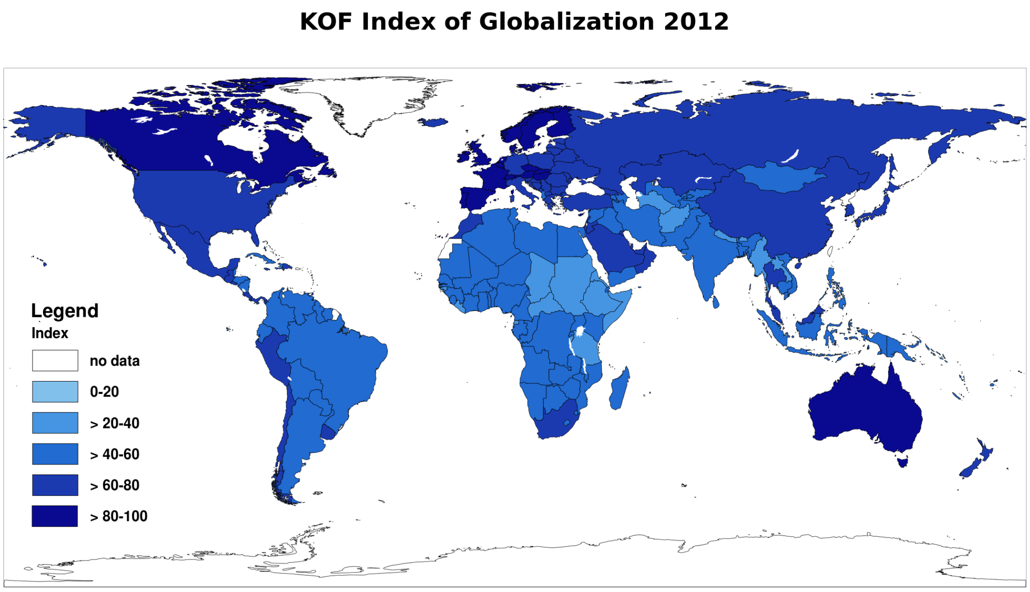

Figure 1.1 *An example of KOF globalisation index estimates for year 2012.*

Source: http://globalization.kof.ethz.ch/

Raab et al. [81] expanded the list of indicators used by Dreher [80] and included a handful of cultural integration indicators, such as the right to education, spread of human rights, gender equality, increase in urbanisation and tertiarisation. The weighing was performed using PCA. Authors noticed increasing cultural convergence and argued its usefulness in sociological context.

Maastricht Globalisation Index (MGI) was suggested by Martens and Zywietz [76] and developed later by Martens and Raza [82], refined afterwards by Figge and Martens [83]. The authors went beyond the dimensions used by Kearney [77], Dreher [80] or Lockwood and Redoano [79] and expanded the indicator set to cover environmental issues and organised violence. Another methodological addition was adjusting the indicators for geographical characteristics by regressing them on logarithm of population and a landlocked dummy. The resulting residuals were further used in index construction by summation giving all the indicators the same weight. Since added new dimensions required certain data which was sparse, index estimates are only available for 3 time periods: 2000, 2008 and 2012, therefore it is not as informative on globalisation dynamics as other indicators.

In 2010 Vujakovic [84] presented a "New Globalisation Index" (NGI) which was evaluated using principal component analysis. The data set used for evaluation consisted of 21 variables, which were similar to the ones used by Kearney [77]. They were assigned by the author into 3 separate groups: economical (trade in goods and services, FDI and portfolio investment statistics, income payments to foreign nationals, trademark and patent applications by non-residents), political (international environmental agreements, international organisation memberships, number of embassies, participation in UN peacekeeping missions) and social (migration, tourism, outbound student mobility, international phone-calls, internet bandwidth and transfers, international trade in newspapers and books) indicators. The data covered 70 countries in the time span of 1995–2005. International trade in goods was adjusted for geographical distance, and several other indicators were adjusted for country size. The first 3 principal components were acquired from this data set and the NGI was calculated for each country as a percentage of variance explained by those 3 principal components. The author was focused on cross-country comparability and transformed the estimated values into ranks, which make the comparison of countries very convenient in a single time point. However, the comparison in rank dynamics might be misleading since the country might increase its international integration but its rank might drop. Also, countries with very similar indices from PCA might be far apart in ranks if there are multiple similar estimates.

The largest merit of Vujakovic [84] approach is capability to find patterns of similarity across very different domains, and the principal of multidimensionality in definition is well incorporated. Nevertheless, the author did not provide the loading estimates from principal component analysis. Therefore, it is indistinguishable if the index attributes positive load for values of indicators expressing greater international integration, it might be that it just measures similarity across countries in certain indicators.

Most of reviewed researchers developed their own indexes of globalisation by adding additional indicators to the ones used by Kearney and/or adjusting the weights. Andersen and Herbertsson [85, 86] made a methodological advance into different direction. The data pool that they used was from economic domain and consisted of nine indicators: freedom to use alternative currencies, freedom of exchange in capital and financial markets, freedom to trade with foreigners, gross private capital flows as a ratio of GDP, export + import of goods and services as a ratio of GDP, factor income received as a ratio of GDP, factor income paid as a ratio of GNP, changes in terms of trade, inflow of direct investment as a ratio of GDP. They performed factor analysis and extracted 2 factors which were used for weighing. Also, they gave the names to the factors in accordance to what indicators they load on the most. The loadings from the first factor were attributed to overall globalisation index based on actual use of international integration and the results from second factor describe the institutional setup for international transactions.

1.1.3. Openness and growth

Many of presented openness measures were applied for studying relationship between openness of international trade policy and economic growth. Using different techniques and various openness measures many authors [47, 52–54, 56, 61, 62, 65, 66, 69, 71] arrived at the same conclusion: greater trade openness is associated with faster growth of economy. In addition to that, the same issue was analysed in the microeconomic context and it was examined whether more productive firms are more likely to become exporters.

One example of sector-level analysis is a study by Nishimizu and Robinson [87]. According to one of the stylized facts in productivity studies, total factor productivity (TFP) is usually apportioned from $\frac{1}{3}$ to $\frac{1}{2}$ of total output growth. They examined the impact of foreign trade policy on TFP for 3 countries — linear regression was build with TFP growth as dependant variable, output growth allocated to export expansion and output growth allocated to import substitution as regressors. The variables were decomposed into their equivalents for 13 manufacturing sectors. The results showed that there are significant differences across industries and countries, and indicated the export-orientated and import-competing industries. They also concluded that foreign trade policy is very important to the growth of TFP. Similar conclusions were drawn by Krueger and Tuncer [88], Nishimizu and Page [89] using the same methodology.

Rodrik [90] pointed out that opening international trade may reduce the rate of catch-up to international productivity levels of import-competing sectors and accelerate it among exporting ones. Furthermore, the firms that are protected from foreign competition may not be willing to modernise their plants. This statement has been confirmed by multiple empirical studies [91–94] stemmed from assumptions on a micro-level, the most important of which is heterogeneity among firms. Tybout [95] pointed out that even though most of these studies suffer from data-related problems, such as data unavailability for small firms, they support the evolutionary principal that market entry is performed by high productivity firms and low productivity firms are more likely to exit. This principal has been demonstrated not only from an importers perspective but continues to apply in the studies analysing export decisions on the micro-level. The findings of Helpman *et al* [96] show that more productive firms choose to serve the foreign markets and the most productive among this group will further choose to serve the overseas market via foreign direct investment.

Dreher [80] addressed the connection between the globalisation and economic growth and found that globalisation indeed promotes growth.

1.1.4. Section conclusions

Studies examining the relation between openness and economic growth produced a handful of measures for international integration. Many of them are based on international trade or restrictions to it. Since the variable of interest is economic growth, most critique aimed at openness measures is based on a notion that those measures should capture only trade policy related variables and not include the effect of other (not controlled) variables which may contribute to economic growth, thus affecting the evaluation of openness effect on economic growth. Despite criticism nearly all (exceptions are Young [97] and Rodrik [98]) researchers found positive relationship between openness and economic growth.

Concerning globalisation measurement, many authors face a difficulty finding an interpretation for the constructed estimates. If the constructed measure is a syndicate index, it is not clear what it measures. Moreover, syndicate measures often rely on the subjective judgement upon selecting indicator weights. The measures that have a clear interpretation are criticised for being too simplistic to capture a multi-dimensional phenomenon so complex as the globalisation.

Facing these difficulties, it is desirable to construct a globalisation measure using a sound statistical technique to diminish the need of researcher's (often subjective) judgement and to have means for quantitative validation of resulting estimates. Another desired feature for this measure is interpretation and clarity of what it measures. Upon this requirement it was decided to adopt the idea that globalisation (and openness) promotes economic growth, which is justified from economic theory and demonstrated empirically by numerous researchers, and measure how much of economic growth is explained by foreign factors. In order grasp the multi-dimensional nature of globalisation an indicator which captures multi-domain economic growth is needed.

1.2. Business cycle indicators

The traditional definition of recession is a decline in real GDP over 2 consecutive quarters [99]. However, the recession dates of U.S. are published by NBER's Business Cycle Dating Committee taking into account other information [100]. Even though the GDP is most popular indicator of economic activity, it reflects only a sum of economic activity and is not informative if the decline occurs in one sector or in overall economy. Therefore alternative indices of economic activity are needed and a lot of authors have contributed to measuring of business cycles.

Indices of economic activity are calculated by all developed countries (e.g. OECD countries) and some major developing economies. The most popular are coincident and leading economic indices which indicate the current and forthcoming business cycle phase. These indices are used to summarise and forecast macroeconomic activity and provide valuable information for policy makers, tax collectors and businesses.

The business cycle fluctuations occur around a long-term growth trend and are most often measured by the growth rate of real GDP. According to Schumpeter [101] recessions are inevitable price to pay for a long-term growth and these economic downturns with innovation force to reorganise production and achieve greater efficiency, lesser costs; and eliminate inefficient non-innovating businesses.

With the attempt to build a methodology to measure the business cycle the economists Burns and Mitchell [102] were the first authors to analyse economic time series to determine if their cyclical turning points lagged, coincided or lead with the business cycle of the economy. A subset of these time series were declared reliable indicators and were monitored by National Bureau in U.S. for indications of broad macroeconomic swings. The crucial criterion selecting these series was location of their turning points and their correspondence to cyclical indicator. According to Burns and Mitchell division these time series were combined into coincident, leading and lagging indices by NBER (National Bureau of Economic Research) economists Shiskin and Moore [103]. The economic indices were constructed by weighted averaging of selected series. The basic principles of this methodology were applied and expanded by Conference Board [104] and OECD [105] by adding additional indicators and refining the weight selection. These methodologies are still in application and are regularly updated. The indicator selection for index construction [106] is based on several criteria: the economic significance, statistical adequacy (in describing the economic process in question), timing of revivals and recessions, matching to historical business cycles, smoothness, promptness of indicator publication.

Chauvet [108] pointed out that constant revisions of Conference Board [104] methodology and weight adjusting is a tedious process. Auerbach [109] develops a more advanced method for variable selection discarding the criterion of turning points and selecting leading indicators using the results from regressing the cyclical indicator on potential leading indicators and their lagged values. Auerbach also refined the computation of weights using the evaluation procedure which maximises the prediction accuracy of coincident indicators using selected leading series. The final leading index is produced by weighted summing as in the original Conference Board methodology. Linear combination of coincident series is also applied by Issler and Vahid (2003) and NBER methodology is heavily relied on in their paper, although their procedure for selecting the indicators was conditioned to correspond to NBER recession index (other authors used comparison with recession index as a quality indicator of their indices).

A new wave of methods for constructing coincident and leading economic indices began with the works of Stock and Watson [110, 111] who applied more sophisticated time series econometrics tools that their predecessors. They took Burns and Mitchell's [102] definition of business cycle, which describes it as expansions occurring at about the same time in many

Table 1.1 *The indicators and the weights used by Conference Board for construction of leading, coincident and lagging economic indices for U.S.*
Source: The Conference Board [107]

Index	Indicator	Weight
Leading	Average weekly hours, manufacturing	0.2781
Leading	Average weekly initial claims for unemployment insurance	0.0334
Leading	Manufacturers' new orders, consumer goods and materials	0.0811
Leading	ISM new order index	0.1651
Leading	Manufacturers' new orders, non-defence capital goods excl. aircraft	0.0356
Leading	Building permits, new private housing units	0.0272
Leading	Stock prices, 500 common stocks	0.0381
Leading	Leading Credit Index	0.0794
Leading	Interest rate spread, 10-year Treasury bonds less federal funds	0.1069
Leading	Avg. consumer expectations for business and economic conditions	0.1551
Coincident	Employees on non-agricultural payrolls	0.2597
Coincident	Personal income less transfer payments	0.1357
Coincident	Industrial production	0.0728
Coincident	Manufacturing and trade sales	0.5318
Lagging	Average duration of unemployment	0.0361
Lagging	Inventories to sales ratio, manufacturing and trade	0.1211
Lagging	Labour cost per unit of output, manufacturing	0.0587
Lagging	Average prime rate	0.2815
Lagging	Commercial and industrial loans	0.0970
Lagging	Consumer instalment credit to personal income ratio	0.2101
Lagging	Consumer price index for services	0.1955

economic activities, followed by similarly general recessions, contractions and revivals. Their statistical framework included a dynamic single factor model (1.1) which was used to evaluate the "unobserved state of the economy", and Kalman [112] filter was applied to estimate its parameters.

$$\begin{cases} \Delta X_t & = \beta + \lambda(L)\Delta C_t + \mu_t; \\ D(L)\mu_t & = \epsilon_t; \\ \phi(L)\Delta C_t & = \delta + \eta_t. \end{cases} \tag{1.1}$$

Here (eq. (1.1)) X_t denotes $n \times 1$ vector of the logarithms of selected coincident series, C_t represents the common unobserved variable, or "index", μ_t is n-dimensional component which represents idiosyncratic movements, ϵ_t and η_t are error terms, β and δ are intercepts, L is lag operator, Δ is difference operator, $\phi(L)$, $\lambda(L)$ and $D(L)$ are respectively scalar, vector and matrix lag polynomials.

The coincident time series included into model(1.1) were selected according to Conference Board recommendations and were from areas of employment, trade, manufacturing and production. The constructed coincident economic index (CEI) reflected co-movements across various economic activities and it is an alternative measure of economic activity to GNP and GDP.

Stock and Watson [110, 111] method of building a leading economic index (LEI) was based on a non-traditional approach: the leading economic index was constructed as a forecast of the coincident index using leading indicators and was evaluated using a simultaneous equation system:

$$\begin{cases} \Delta C_t & = \mu_C + \lambda_{CC}(L)\Delta C_{t-1} + \lambda_{CY}(L)Y_{t-1} + \nu_{Ct}; \\ Y_t & = \mu_Y + \lambda_{YC}(L)\Delta C_{t-1} + \lambda_{YY}(L)Y_{t-1} + \nu_{Yt}. \end{cases} \tag{1.2}$$

Here (eq. (1.2)) Y_t is a vector of leading series and (ν_{Ct}, ν_{Yt}) are serially uncorrelated error terms.

In comparison to previous work, this LEI evaluation method has all advantages of econometric methods to check if the model is adequate and the variables used are statistically significant. Stock and Watson also proposed a new recession index which is interpreted as a probability that the economy will be in a recession six months hence.

The methodology of Stock and Watson [110, 111] was enthusiastically accepted by other researchers and different expansions and alterations to it were proposed. Diebold and Rudebusch [113] suggested a dynamic factor model with regime switching which was proved to perform very similarly to the Department of Commerce methods and Stock and Watson methods. Moreover it did bring the upside of regime switching methods: improved the forecast performance and the ability to track switches in optimal decision rules (e.g. in consumption or investment) which may occur with regime change. McGuckin et al. [114] suggested incorporating financial information and forecasts of real variables into construction and proved this inclusion to be useful and acquired the increase in accuracy. Similar idea was used by Estrella and Mishkin [115] and their results also indicated that financial data has very informative leading indicators.

Mariano and Kurosawa [116] offered the adaptation of Stock and Watson methodology for monthly coincident index evaluation for countries which measure the GDP in quarterly terms. This method has a certain appeal as their coincident index has a strong relation to latent monthly real GDP and is therefore easier for interpretation.

Evaluation methods were also a subject of new suggestions. An alternative method for evaluating the dynamic single-factor model (other than Kalman filtering) is the Bayesian approach applied by Otrok and Whiteman [117]. The advantage of this method is the possibility to extract not only the mean, but the whole distribution of the latent factor. Another alternative was evaluating the factor model in the frequency domain, which was suggested by Forni et al. [118]. This method provides more flexibility on assumptions in comparison to original Stock and Watson [110, 111] approach.

Section conclusions

Despite the overall popularity of GDP as an indicator of economic activity, alternative measures are preferred when measuring the business cycle. The main argument for that is the requirement that business cycle indicator should reflect the fluctuations that are common across different sectors of economy and the sum of all economic activity given by GDP is not sufficient to indicate that. In order to attain such measure the main task is extracting a common pattern from multiple coincident indicators, and factor models suit this purpose very well. A variety of factor model versions for coincident economic index evaluation were proposed based on pioneering works of Stock and Watson [110, 111] and their methodology is heavily relied on in the empirical part of this dissertation, i.e. constructing the coincident economic index using a dynamic single-factor model.

1.3. Factor models

1.3.1. Principal components and factor analysis

The roots of factor modelling lie within the principal component analysis. It is a statistical technique frequently used to reduce the dimensionality [119]. It was first introduced by Pearson in 1901 [120] and developed independently by Hotelling in 1933 [121]. It is based on the idea of orthonormal decomposition.

Definition of principal components Suppose that $\mathbf{x}$ is a vector of p variables with a covariance matrix $\mathbf{\Sigma}$. The linear expression $\boldsymbol{\alpha}_1^\top \mathbf{x}$ of the elements of $\mathbf{x}$ where $\boldsymbol{\alpha}_1 = (\alpha_{11}, \alpha_{12}, ..., \alpha_{1p})$ having maximum variance is called the first principal component of $\mathbf{x}$. The second principal component $\boldsymbol{\alpha}_2^\top \mathbf{x}$ has a maximum variance under the constraint that $\boldsymbol{\alpha}_1^\top \mathbf{x}$ is uncorrelated with $\boldsymbol{\alpha}_2^\top \mathbf{x}$. Up to p principal components could be found.

One of the core properties of principal components which is very useful in algebraic computations is that kth principal component can be expressed by $z_k = \boldsymbol{\alpha}_k^\top \mathbf{x}$ where $\boldsymbol{\alpha}_k$ is an eigenvector of $\mathbf{\Sigma}$ corresponding to its kth largest eigenvalue λ_k. Furthermore, if $\boldsymbol{\alpha}_k$ is chosen to have unit length $\boldsymbol{\alpha}_k^\top \boldsymbol{\alpha}_k = 1$ then $\mathrm{var}(z_k) = \lambda_k$ [119].

Factor analysis model

If we have p observed random variables $x_1, x_2, ..., x_p$ they can be expressed, except for an error term, as linear functions of $m(< p)$ hypothetical (random) variables or common factors $f_1, f_2, ..., f_m$:

$$
\begin{aligned}
x_1 =& \lambda_{11} f_1 + \lambda_{12} f_2 + ... + \lambda_{1m} f_m + e_1 \\
x_2 =& \lambda_{21} f_1 + \lambda_{22} f_2 + ... + \lambda_{2m} f_m + e_2 \\
&\vdots \\
x_p =& \lambda_{p1} f_1 + \lambda_{p2} f_2 + ... + \lambda_{pm} f_m + e_p
\end{aligned}
\tag{1.3}
$$

The expression (1.3) or matrix version (1.4)

$$
\mathbf{x} = \mathbf{\Lambda} \mathbf{f} + \mathbf{e} \tag{1.4}
$$

are pretty general forms used in factor analysis. The following are most common assumptions within the framework of factor analysis:

1. $\mathbf{E}[\mathbf{e}] = 0$
2. $\mathbf{E}[\mathbf{f}] = 0$
3. $\mathbf{E}[\mathbf{x}] = 0$
4. $\mathbf{E}[\mathbf{e}\mathbf{e}^\top] = \mathbf{\Psi}$ (diagonal)
5. $\mathbf{E}[\mathbf{f}\mathbf{e}^\top] = \mathbf{0}$ (matrix of zeroes)
6. $\mathbf{E}[\mathbf{f}\mathbf{f}^\top] = \boldsymbol{I}_m$ (identity matrix)

The equations (1.3) describe a model which is the main difference between principal component analysis and factor analysis. However the first m principal components are frequently used as initial values or approximation of the factors in (1.3).

The equations (1.3) look like generic linear regressions, but in this case the both terms $\mathbf{\Lambda}$ and $\mathbf{f}$ are unknown, therefore the best fitting solution is not unique. The covariance of both sides is

$$
\mathbf{\Sigma} = \mathbf{\Lambda} \mathbf{\Lambda}' + \mathbf{\Psi} \tag{1.5}
$$

If $(\mathbf{\Lambda}, \mathbf{\Psi})$ is a solution and $\mathbf{T}$ is a quasi-orthogonal matrix, then $(\mathbf{\Lambda T}, \mathbf{\Psi})$ is also the solution since $(\mathbf{\Lambda T})(\mathbf{\Lambda T})' = \mathbf{\Lambda T T' \Lambda'} = \mathbf{\Lambda \Lambda'}$. For this, the estimation is performed in 2 stages: first, $\mathbf{\Lambda}$ and $\mathbf{\Psi}$ are found after placing some restrictions on $\mathbf{\Lambda}$, then $\mathbf{f}$ is calculated. Having an initial solution $\mathbf{\Lambda}$ other solutions can be found using so called *rotation* procedure, i.e. multiplying $\mathbf{\Lambda}$ by an orthogonal matrix. There are multiple criteria for selecting a "good" orthogonal matrix for rotation. Usually the requirement is for a final solution to have a lot of elements either "close to zero" or "far from zero" so that it would be easier to come up with interpretation for the results.

1.3.2. Factor models for time series

Factor models for time series evaluate factors that are also of time series class. Dynamic factor models require assumptions on evolution of factors that are defined in a system of equations. The most popular assumption is of autoregressive form and the dynamic factor model can be expressed as a Gaussian linear state-space model, also called dynamic linear model (DLM) [122]:

$$\mathbf{x_t} = \mathbf{\Lambda F_t} + \mathbf{e_t} \tag{1.6}$$

$$\mathbf{F_t} = \mathbf{\Phi F_{t-1}} + \mathbf{u_t} \tag{1.7}$$

Here $\mathbf{x_t}$ is n-dimensional time series, $\mathbf{\Lambda}$ is $n \times m$ matrix, containing unknown factor loadings, $\mathbf{F_t}$ is a vector of m factors. Is is generally assumed that $\mathbf{e_t} \sim \mathcal{N}(\mathbf{0}, \mathbf{W})$, $\mathbf{u_t} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$, $\mathbf{F_0} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$ and that $\mathbf{e_t}$, $\mathbf{u_t}$ and $\mathbf{F_0}$ are independent of each other but interdependence assumption is not strictly necessary [123]. The first equation in this system is called *observation equation* and the second one is *state equation*. The unknown parameters of dynamic linear model are $\mathbf{\Lambda}$, $\mathbf{\Phi}$, $\mathbf{W}$, $\mathbf{V}$, $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$, and they are called hyperparameters.

The assumptions of state-space model are:

1. $\{\mathbf{F}_t, t = 0, 1, \ldots\}$ is a Markov chain, that is $\mathbf{F}_t$ depends on the past values of $\{\mathbf{F}_0, \mathbf{F}_1, \ldots\}$ only through $\mathbf{F}_{t-1}$. Therefore the probability law of the process is specified by setting the initial density $p_0(\mathbf{F}_0)$ and transition densities $p(\mathbf{F}_t|\mathbf{F}_{t-1})$.

2. Conditionally on $\{\mathbf{F}_t, t = 0, 1, \ldots\}$ the $\mathbf{x_t}$ are independent of each other, and $\mathbf{x_t}$ are dependent on $\mathbf{F}_t$ only. It follows that for any $n \geq 1$, $(\mathbf{x_1}, \ldots, \mathbf{x_n})|\mathbf{F}_1, \ldots, \mathbf{F}_n$ have a joint conditional density $\prod_{i=1}^{n} p(\mathbf{x_i}|\mathbf{F}_i)$

Even though equation (1.7) defines AR(1) process, higher order AR processes can be defined by adding lagged factors into $\mathbf{F_t}$ and imposing certain restrictions. For example AR(2) process could be defined by specifying equation (1.7) like this:

$$\begin{bmatrix} \mathbf{F_t} \\ \mathbf{F_{t-1}} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi_1} & \mathbf{\Phi_2} \\ \boldsymbol{I}_m & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{F_{t-1}} \\ \mathbf{F_{t-2}} \end{bmatrix} + \begin{bmatrix} \mathbf{u_t} \\ \mathbf{0} \end{bmatrix} \tag{1.8}$$

Conditional on hyperparameters the variance of $\mathbf{x_t}$ is given by

$$\mathrm{var}(\mathbf{x_t}) = \mathbf{\Lambda}\mathrm{var}(\mathbf{F_t})\mathbf{\Lambda}' \tag{1.9}$$

The joint log-likelihood of the observed time series and common factors is:

$$\begin{aligned}
\log L(\mathbf{x_1}, \ldots, \mathbf{x_T}, \mathbf{F_0}, \ldots, \mathbf{F_T}) = &-\frac{1}{2}\log|\mathbf{\Sigma}| - \frac{1}{2}(\mathbf{F_0} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{F_0} - \boldsymbol{\mu}) \\
&- \frac{T}{2}\log|\mathbf{V}| \\
&- \frac{1}{2}\sum_{t=1}^{T}(\mathbf{F_t} - \mathbf{\Phi}\mathbf{F_{t-1}})'\mathbf{V}^{-1}(\mathbf{F_t} - \mathbf{\Phi}\mathbf{F_{t-1}}) \\
&- \frac{T}{2}\log|\mathbf{W}| + const. \\
&- \frac{T}{2}\sum_{t=1}^{T}(\mathbf{x_t} - \mathbf{\Lambda}\mathbf{F_t})'\mathbf{W}^{-1}(\mathbf{x_t} - \mathbf{\Lambda}\mathbf{F_t})
\end{aligned} \tag{1.10}$$

The expression includes unknown factor components, therefore maximum likelihood method cannot be applied directly on this expression. The evaluation of parameters might be performed by EM algorithm or Kalman filtering and smoothing, which are explained step-by-step by Zuur *et al* [123]. Another alternative for parameter evaluation is using Bayesian methods.

1.3.3. Bayesian methodology

Complicated likelihood functions can be rewritten using the notion of conditional independence. This method is named after Thomas Bayes who came up with a formula for conditional probability:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1.11}$$

The assumption on distribution of parameter set θ is called prior. The basic principal of conditional independence is assuming that observations $X_1, X_2, \ldots X_n$ are independent conditionally on a set of parameters θ with a density $\pi(\theta)$. The posterior distribution of θ is dependent on data and using Bayes formula can be expressed [122] the following way:

$$\pi(\theta|x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n|\theta)\pi(\theta)}{m(x_1, \ldots, x_n)} \propto \prod_{t=1}^{n} f(y_t|\theta)\pi(\theta) \tag{1.12}$$

The marginal density $m(x_1, \ldots, x_n)$ (1.13) does not depend on θ and has a normalising role.

$$f(x_1, x_2, \ldots, x_n) = \int f(x_1, x_2, \ldots, x_n|\theta)\pi(\theta)\mathrm{d}\theta \tag{1.13}$$

Since Bayesian methodology is most often used when other, simpler, methods are infeasible, it is likely that the posterior distribution of the parameters is analytically intractable. In order to overcome these limitations it is usually resorted to simulation methods. Monte Carlo methods based on simulating random variables from a Markov chain, called Markov chain Monte Carlo (MCMC) methods, are nowadays the standard way of evaluating posterior distribution and parameters required by Bayesian methodology [122].

Markov chain Monte Carlo

MCMC method was firstly introduced by Metropolis *et al* [124] to solve a statistical physics problem, and was later generalised by Hastings [125] with a focus on statistical problems. A version of MCMC which is frequently referred to as Gibbs sampler was developed by Geman and Geman [126] and proved to be very useful for general Bayesian computation [127].

For a Markov chain $\{\theta_t\}_{t>0}$ meeting certain requirements with distribution π, it can be shown that for every initial value θ_1 the distribution of θ_t tends to π as t increases to infinity. A more formal formulation for this result is as follows [128].

Let $\{X_0, X_1, \ldots, X_t, \ldots\}, X_t \in E \subseteq \mathbb{R}^n$ be a Markov chain with transition kernel $K : E \times E \to \mathbb{R}^+$ such that with respect to σ-finite measure ν on the Borel σ-field of $\mathbb{R}^n$ for ν-measurable A,

$$P(X_t \in A | X_{t-1} = x) = \int_A K(x, y) \mathrm{d}\nu(y) + r(x) \mathbb{1}_{\{x \in A\}}$$

where

$$r(x) = 1 - \int_E K(x, y) \mathrm{d}\nu(y)$$

K is called π-irreducible if,

$$\forall x \in E, \pi(A) > 0 \Rightarrow \exists t \geq 0 : P(X_t \in A | X_0 = x) > 0.$$

K is called aperiodic if there does not exist a partition $E = (B_0, \ldots, B_{r-1})$ for some $r \geq 2$, such that $P(X_t \in B_{t\mathrm{mod}(r)} | X_0 = x_0 \in B_0) = 1, \forall t$.

Theorem 1 *If K is π-irreducible and aperiodic then for all*
$x \in D = \{x \in E, \pi(x) > 0\}$*:*

1. *K converges in ν-measure to π, $t \to \infty$*

2. *for real-valued π-integrable f,*

$$\frac{f(X_1) + f(X_2) + \ldots + f(X_t)}{t} \to \int_E f(x)\pi(x)\mathrm{d}\nu(x)$$

almost surely as $t \to \infty$

MCMC method could also be applied for evaluation of the unknown hyperparameters of a state space model, which is usually the case. For that is it usually assumed that unknown parameters depend on a set of variables ψ. Using Bayesian idea, the ψ is assumed to be a random vector, and the assumptions on state-space models are assumed to hold conditionally on ψ. Prior knowledge about ψ is expressed through a probability law $\pi(\psi)$. Thus, for any $n \geq 1$, it is assumed that

$$(\mathbf{F}_0, \mathbf{F}_1, \ldots, \mathbf{F}_t, \mathbf{x}_1, \ldots, \mathbf{x}_t, \psi) \sim \pi(\mathbf{F}_0|\psi)p(\psi)\prod_{t=1}^{n} f(\mathbf{x}_t|\mathbf{F}_t, \psi)\pi(\mathbf{F}_t|\mathbf{F}_{t-1}, \psi)$$

Given the data $\mathcal{D}_t$ the unknown states and parameters might be evaluated by computing the posterior distribution:

$$\pi(\mathbf{F}_s, \psi|\mathcal{D}_t) = \pi(\mathbf{F}|\psi, \mathcal{D}_t)\pi(\psi|\mathcal{D}_t)$$

If we denote $(\mathbf{F}_0, \mathbf{F}_2, \ldots, \mathbf{F}_t)$ as $\mathbf{F}_{0:t}$, then the joint distribution of interest is:

$$\pi(\mathbf{F}_{0:t}, \psi|\mathcal{D}_t) = \pi(\mathbf{F}_{0:t}|\psi, \mathcal{D}_t)\pi(\psi|\mathcal{D}_t) \tag{1.14}$$

Gibbs sampling algorithms can be used for approximating the joint posterior $\pi(\mathbf{F}_{0:t}, \psi|\mathcal{D}_t)$. It requires to iteratively simulate from conditional distributions $\pi(\mathbf{F}_{0:t}|\psi, \mathcal{D}_t)$ and $\pi(\psi|\mathcal{D}_t)$ in relation to (1.14).

The customary MCMC approach to analyse the posterior distribution $\pi(\mathbf{F}_{0:t}, \psi|\mathcal{D}_t)$ is to generate a dependent sample from it and evaluate posterior summaries from the simulated sample [122]. The simulated sample from the posterior can in turn be used as input to generate a sample from the predictive distribution of states and observables. This approach solves the filtering, smoothing and forecasting problems for a DLM with unknown parameters.

1.3.4. Application

The factor models and dynamic linear models are frequently used in prediction problems. The prediction using dynamic linear models (DLM) is straightforward using transition equation for state forecast and then multiplying it by loading parameters, i.e. setting $\mathbf{u}_{t+1}$ and $\mathbf{e}_{t+1}$ to zero.

Another use of factors in prediction was employing them to summarise information using a large number of predictors. This idea was first presented by Sargent and Sims [129] with an argument that traditional models stemming from Kaynes economic theory tend to have over-identified equations that do not necessarily reveal true statistical relationship between variables. Stock and Watson adopted the Sargent and Sims' idea [129] and proposed their version of forecasting using diffusion indexes [130,131]. They also proved the forecasts to be consistent and asymptotically efficient.

The model by Stock and Watson consists of the following equations:

$$X_{it} = \lambda_i(L) f_t + e_{it} \tag{1.15}$$

$$y_{t+1} = \beta(L) f_t + \epsilon_{t+1} \tag{1.16}$$

Here y_t denotes the time series to be forecast, X_t is N-dimensional multiple time series of predictors. It is assumed that (X_t, y_{t+1}) admit a dynamic factor model representation with r common factors f_t. e_t is idiosyncratic disturbance, $\lambda_i(L)$ and $\beta(L)$ are lag polynomials, f_t and e_t are assumed to be stationary processes of zero mean, so X_t and y_t are deviations from their means.

Additional methodological findings are associated with this method. Bai and Ng [132] proposed a criterion for selecting an optimal number of factors for a forecasting task using Stock and Watson [130] approach. Forni *et al* [118] proposed a modification of this method by constructing a richer dynamic structure and factor evaluation in frequency domain. Bai and Ng [133] demonstrated that using a data set of targeted predictors (i.e. selecting a subset of indicators based on certain criteria) gives a significant boost in forecasting accuracy.

1.3.5. Hierarchical dynamic factor models

Hierarchical linear models were introduced by Lindley and Smith [134] and were extended to dynamic hierarchical linear models by Gamerman and Migon [135]. Those models could be specified to have different number of levels. A three-level hierarchical dynamic linear model is constituted of 3 equations:

$$Y_t = \Lambda_G G_t + v_t \tag{1.17}$$

$$G_t = \Lambda_F F_t + e_t \tag{1.18}$$

$$F_t = \phi F_{t-1} + w_t \tag{1.19}$$

The disturbances $v_t \sim \mathcal{N}(0, V_Y), e_t \sim \mathcal{N}(0, V_G), w_t \sim \mathcal{N}(0, W)$ are assumed independent and the matrices Λ_F and Λ_G are of full rank. More general forms of specification are possible, such as allowing time-varying parameters Λ_G, Λ_F or V_Y, V_G, W.

Dynamic hierarchical factor models were offered by Moench *et al* [15] to use in large datasets and construct the model according to the data structure, thus having a direction for interpretation. Their approach was to organise data into blocks and have separate factors evaluated for each block and model the block-level factors at higher level equations. They showed that dynamic hierarchical factor models (DHFM) are useful to monitor complex data structures and assessing the within and between block variations. For example, DHFM has been used to evaluate how much housing prices are dependent on regional variables by imposing a block structure using geographical division [16]. The process of organising the data into blocks could also help to improve balance since some blocks could be significantly larger that others.

1.3.6. Section conclusions

Factor models and dynamic linear models are powerful statistical tools which are used in common pattern detection among different indicators. These models also have gained popularity in prediction problems from a large number of predictors. The evaluation of parameters of these models is cumbersome since the complex dynamic structure requires advanced evaluation methods, plus the number of unknown parameters is large and some restrictions (or assumptions) are required in order to process the calculations.

Dynamic hierarchical factor models allow to impose a structure on the model that corresponds to the data structure and evaluate correlated factors if they are in different blocks. This approach is convenient to come up with interpretation for acquired factors opposite to factor analysis models which leave the problem of finding an adequate rotation and interpretation for the researchers' insight and intuition. This particular feature of DHFMs is very appealing if researcher *a priori* knows what kind of interpretation she is aiming at. Therefore such approach would be suited to the task of assessing the domestic and foreign factors in order to evaluate their effect on economic growth.

2. Research framework

In the intent to develop a new measure which reveals the extent and dynamics of foreign impact conditioned by the process of globalisation, a research framework was built based on dynamic hierarchical factor models. In order to capture the effect of supranational variables a small open economy was a plausible selection, therefore Lithuania was chosen for it meets this criterion and is familiar to the author since it is my home-country.

The review of other authors' research implies that openness (or, is some cases, globalisation) positively affects economic growth. This result is taken as given with the intent to evaluate the significance of this effect and acquire a new measure of how globalised a focal economy which is based on the magnitude of foreign effect on the growth of economy. Therefore prediction based measurement could indicate country's sensitivity to global shocks and reveal how much focal country's economy is intertwined with global economy. This way the proposed measure is clear about what it measures opposite to syndicate measures. Another point of interest is to examine the dynamics of this effect and evaluate if the rate of globalisation is increasing over time.

The results of other authors also stress the multi-dimensionality of the process of globalisation therefore it is desirable to assess the effect on the growth of economy using a measure of economic activity which reflects that. The coincident economic index has this exact interpretation and therefore the first step was coincident economic index evaluation.

2.1. The coincident economic index

According to Stock and Watson [110], the coincident economic index (or CEI) reflects the "unobserved state of the economy" and is coincident with the business cycle, which consists of expansions and contractions occurring at the same time in many economic activities and commonly refers to co-movements in different forms of economic activity.

2.1.1. Potential coincident variables

Variables on the subject of output, employment and retail were considered to include in the dynamic single-factor model following the Stock and Watson methodology [110]. Variables of those subjects are commonly used by many methodologies for construction of the coincident economic index (OECD methodology [105], Conference Board methodology [104]).

The initial list of variables that were considered including in the Stock-Watson dynamic single factor model consists of employee hours in non-agricultural establishments, wholesale-retail, income from manufacturing, index of employment in the construction sector. Several variables were considered to take from each category (i.e. output, employment and retail). Selection was based on availability and their relationship to the business cycle. Since the employment seemed to be lagging behind the business cycle, it was left out the model. Another variable that was decided to include in the list is the index of real estate prices. The motivation for doing this is that the Lithuanian economy was severely affected by the real estate bubble and the rapid growth of the construction sector, which is fairly well described by housing prices.

The final list of the variables selected for the dynamic single factor model is [136]:

- IM – Turnover of manufacturing
- RE – Real estate price index
- WT – Turnover index of wholesale trade
- IP – Index of production

These series are quarterly seasonally adjusted[1] data [2] covering period from 1998 1st quarter to 2013 3rd quarter. Since RE series started at the 4th quarter of the year 1998, the values of first three quarters were extrapolated backwards using Holt-Winters procedure. The initial data analysis showed that these four series are $I(1)$ processes, but they are not cointegrated[3]. The selected variables are plotted in figure 2.1.

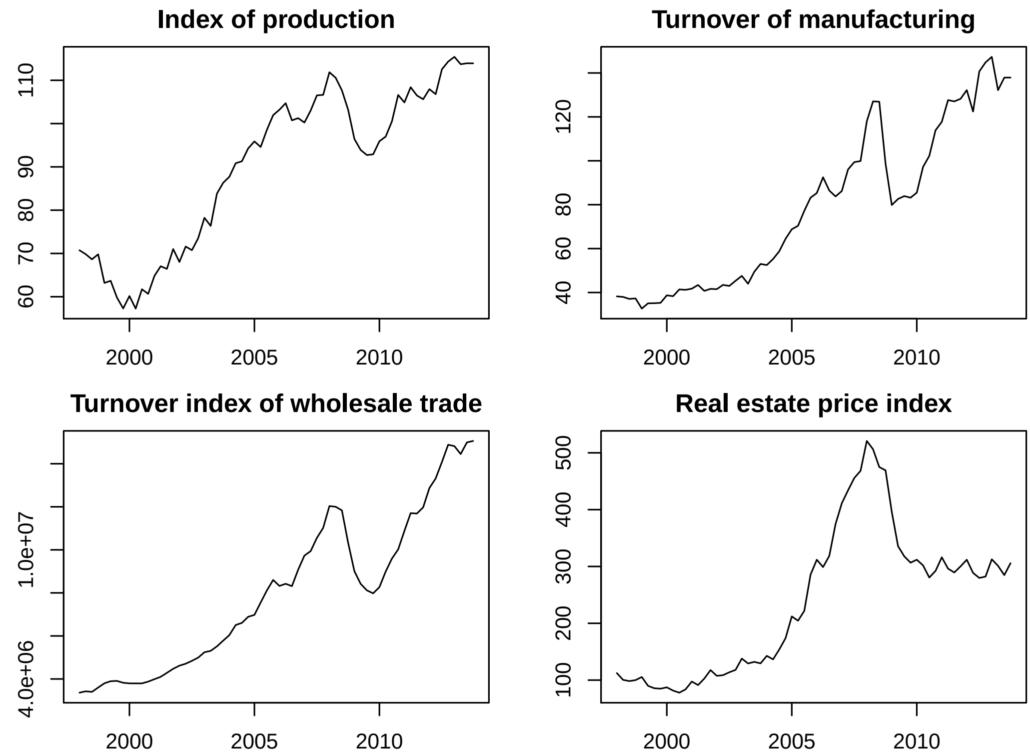

Figure 2.1 *Variables used for coincident index construction*

[1]The seasonal adjustment procedure used was X-13ARIMA-SEATS developed by US Census Bureau (http://www.census.gov/srd/www/x13as/)

[2]*IM*, *WT*, *IP* series were acquired from Statistics Lithuania. The source of *RE* series is State Enterprise Centre of Registers

[3]Dickey-Fuller test failed to reject the null hypothesis about unit root existence and Johansen test did not provide evidence about cointegration

2.1.2. Evaluation procedure

A dynamic single factor model was built following Stock and Watson [110]. The coincident economic index is a transformation of the estimate of a single factor – "the unobserved state of the economy". The structure of the constructed model is given in equations (2.1), (2.2), (2.3), (2.4).

$$\Delta X_t = \beta + \gamma(L)F_t + \mu_t, \tag{2.1}$$

$$D(L)\mu_t = \varepsilon_t, \tag{2.2}$$

$$\psi(L)F_t = \delta + \eta_t, \tag{2.3}$$

$$\Delta C_t = a + bF_t. \tag{2.4}$$

Here X is a vector of logarithms of coincident variables IM, RE, WT and IP. F_t is a factor, describing the unobserved state of the economy at time t. The functions $\psi(L)$, $\gamma(L)$ and $D(L)$ are respectively scalar, vector and matrix lag polynomials. The error term μ_t is serially correlated and its dynamics is described in equation (2.2). C_t is the coincident economic index. Error terms (ε_t, η_t) are assumed to be serially uncorrelated with the zero mean and diagonal variance matrix Σ. Since F_t has a zero mean and unit variance (step 3 in evaluation algorithm), a and b are the de-normalisation parameters.

Equations (2.1), (2.2) and (2.3) form a state-space model. Its parameters and the "unobserved state of the economy" are evaluated using the Kalman filter.

The evaluation is performed in this order:

1. Each economic variable from vector X is first-differenced: $\Delta X_t = X_t - X_{t-1}$.

2. Each series of differences ΔX_t is normalized by subtracting its mean and dividing by its standard deviation. Since ΔX_t has a zero mean

there is no need to evaluate parameters β (in equation (2.1)) and δ (in equation (2.3)) as they are equal to 0. The normalisation procedure was performed so that each series were of the same importance.

3. After evaluating the parameters of the state-space model with Kalman filter, a new time series F_t is acquired. This has a zero mean and unit variance, because ΔX_t is normalized.

4. F_t is de-normalized (equation (2.4)) and the coincident economic index C_t is constructed:

$$C_t = \begin{cases} c, & t = 0; \\ c + \sum_{i=1}^{t} \Delta C_i, & t = 1, 2, ...T. \end{cases} \tag{2.5}$$

Green and Beckman [137] evaluated parameters a (the trend parameter) and b (the variance around that trend) as a weighted average of the trends of the coincident series, selected into the model, with weights proportional to the contributions of the indicators in the Kalman filter. An alternate method of Crone and Clayton-Matthews [138] sets a to be equal to the GDP growth trend, and the b parameter is evaluated in the same way as Green and Beckman [137]. Since neither of these methods provided desirable results for the Lithuanian economy, a new method was in need. This is based on minimizing the sum of squares: $\sum_{i=t}^{T}(C_t - GDP_t)^2$ (the OLS method was selected expecting to get the same periods of expansion and contraction for the coincident index and Lithuanian real GDP). This procedure can be shown combining equations (2.4) and (2.5):

$$C_t = \sum_{i=1}^{t} \Delta C_i + c = \sum_{i=1}^{t} (a + bF_i) + c = ta + b\sum_{i=1}^{t} F_i + c.$$

This kind of equation can be rewritten in the form of a linear regression which is estimated using OLS:

$$GDP_t = at + b\sum_{i=1}^{t} F_i + c + \varepsilon_t.$$

It is worth mentioning that CEI is not an estimate of GDP (although it might look like one). CEI as well as GDP are both indicators of macroeconomic activity each of them having their own peculiarities.

2.1.3. Estimates

The following measurement equations were evaluated:

$$\Delta IM_t = \lambda_{IM} F_t + \varepsilon_t^{IM}, \tag{2.6}$$

$$\Delta RE_t = \lambda_{RE} F_t + \varepsilon_t^{RE}, \tag{2.7}$$

$$\Delta WT_t = \lambda_{WT} F_t + \varepsilon_t^{WT}, \tag{2.8}$$

$$\Delta IP_t = \lambda_{IP} F_t + \mu_t^{IP} \tag{2.9}$$

The transition equations were:

$$F_t = \psi F_{t-1} + \varepsilon_t^F, \tag{2.10}$$

$$\mu_t^{IP} = d^{IP} \mu_{t-1}^{IP} + \varepsilon_t^{IP}, \tag{2.11}$$

The maximum likelihood estimates of parameters of these equations are listed in the table 2.1.

The variances σ_{RE}^2, σ_{IM}^2, σ_{WT}^2, σ_{IP}^2 are of error terms ε^{RE}, ε^{IM}, ε^{WT}, ε^{IP} respectively.

The constructed coincident economic index (CEI) and scaled GDP are plotted in figure 2.2. It can be indicated from the graph that the CEI reflects the state of economy in a very similar way as the GDP.

2.2. Economic growth prediction on CEI

2.2.1. The leading indicators

According to Stock and Watson [110] methodology the leading index is constructed as a forecast of the coincident economic index (CEI) growth. They use the leading indicators as predictors to build the leading economic

Table 2.1 *Maximum likelihood estimates of equations'* (2.6), (2.7), (2.8), (2.9), (2.10), (2.11) *parameters*

Coefficient	Estimate	St. error	z-statistics	p-value
λ_{IM}	0.070228	0.015191	4.623089	0.0000
λ_{RE}	0.443925	0.147460	3.010487	0.0026
λ_{WT}	0.534319	0.136212	3.922711	0.0001
λ_{IP}	0.382733	0.117485	3.257706	0.0011
ψ	0.622137	0.173196	3.592102	0.0003
d^{IP}	-0.591497	0.108954	-5.428888	0.0000
σ^2_{RE}	0.6611365	0.353168	1.872017	0.0612
σ^2_{IM}	0.0058533	0.000403	14.52534	0.0000
σ^2_{WT}	0.5161726	0.1621033	3.184220	0.0015
σ^2_{IP}	0.4292031	0.1882456	2.284726	0.0223

index. In this study a much larger number of potential predictors is considered therefore a linear regression would not be feasible since there would be too many parameters to evaluate. The intent is to use the linear forecast method by Stock and Watson [131] which was originally developed for macroeconomic forecasting using diffusion indexes. This way I am going to use factors acquired from leading indicators rather than indicators themselves. The constructed prediction equation is of the form of (2.12).

$$\Delta C_{t+2} = \alpha_1(L)G_{1,t} + \alpha_2(L)G_{2,t} + \beta(L)\Delta C_t + \varepsilon_t \tag{2.12}$$

Here ΔC_{t+2} is future growth of CEI, $G_{1,t}$ and $G_{2,t}$ are factors acquired from domestic and foreign indicators, $\alpha_1(L)$, $\alpha_2(L)$, $\beta(L)$ are lag polynomials. The prediction horizon was selected to be 2 quarters because the publication of most macroeconomic indicators is usually performed about a month after a quarter ends, so in order to have prediction on the future

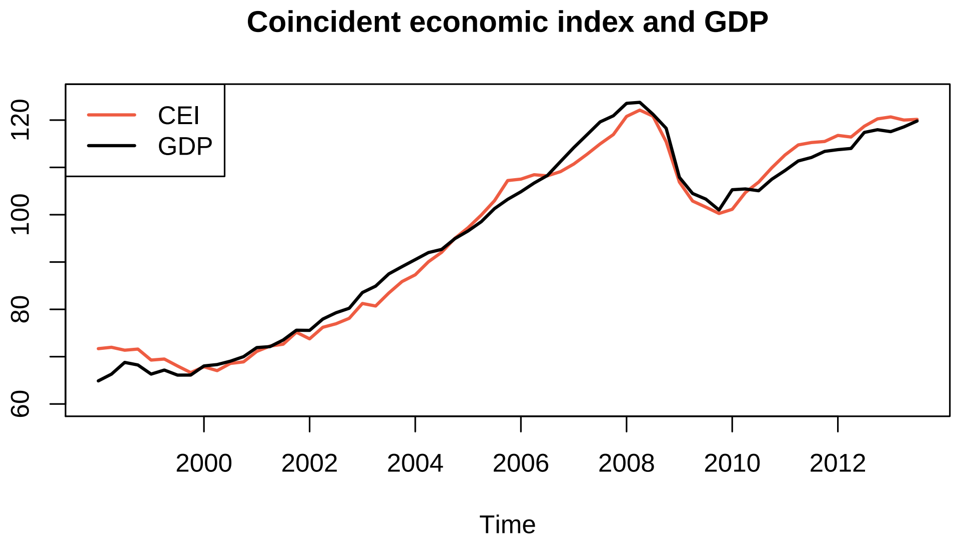

Figure 2.2 *The constructed CEI and its comparison to GDP*

economic growth it was decided to have a bigger prediction span.

The initial domestic data set consisted of 283 time series of most Lithuanian quarterly economic indicators starting at least at 1998 (from the sectors of manufacturing and production, labour, investment, international trade, retailing, public sector, business statistics, construction, transportation and agriculture). The initial supranational data set consisted of 1707 time series which geographically covered Lithuania's top 20 international trade partners[4], groups of countries such as EU, OECD, Euro area and a few largest economies on account that they might have influence to Lithuania through their global presence, such as USA and Japan. The economic indicators were from areas of national accounts, labour statistics, real effective exchange rate, saving and lending. The series were used in real terms where applicable, they were also seasonally adjusted[5] and transformed to be sta-

[4]The number 20 was selected on the account that Lithuania's top 20 trade partners on average cover 90% of exports and 92% of imports and the rest of partners were discarded as having insignificant influence

[5]The seasonal adjustment procedure used was X-13ARIMA-SEATS developed by US Census Bureau (http://www.census.gov/srd/www/x13as/)

tionary.

In order to achieve a straightforward interpretation I am aiming for 1 domestic leading factor and 1 foreign leading factor. Therefore, it is important to use the time series that carry the most information about future growth of economy. Bai and Ng [133] showed that using targeted predictors, i.e. a selected subset from initial dataset, gives a better forecasting accuracy with the same number of factors than using the factors extracted from full data set. For this reason the procedure of leading indicators selection was applied. It is noteworthy that the selection is based on statistical properties of indicators therefore it slightly deviates from the leading indicator definition as used in OECD [105] methodology; the definition used in this study is less restrictive.

The first stage of selecting the leading series was of hard threshold based on two criteria:

1. Granger causality (pairwise testing for lag depth 2 with significance level $\alpha = 0.05$)

2. Correlation between series $\Delta X_{i,(t-l)}$ and coincident index ΔC_t should be greater with lags $l > 0$

Only the series that met both criteria were included into the following stages of modelling. After the first selection stage was completed the data set which consisted of 4 domestic and 15 foreign indicators included several collinear time series, e.g. 6 time series on labour productivity in different European countries and the EU were selected and it was very likely that they carry very similar information. Even though the collinearity does not cause technical problems for factor model evaluation, it can cause a certain imbalance since the factor might hinge to the series that have multiple collinear counterparts.

Detecting and removing collinear series

In order to identify the collinear time series, the data was scaled and euclidean distance (2.13) was calculated between each pair of time series.

$$d(\mathbf{x}_i, \mathbf{x}_j) = ||\mathbf{x}_i - \mathbf{x}_j||_{l_2} = \Big(\sum_t |x_{i,t} - x_{j,t}|^2\Big)^{\frac{1}{2}} \tag{2.13}$$

Afterwards the hierarchical clustering was performed. Initially each series was assigned its own cluster. Next, the most similar series were joined together into a cluster. At each stage the distances were updated using a complete link dissimilarity update formula:

$$d_{i\cup j,k} = \max(d_{i,k}, d_{j,k}) \tag{2.14}$$

The results were combined into cluster dendrogram (Fig. 2.3).

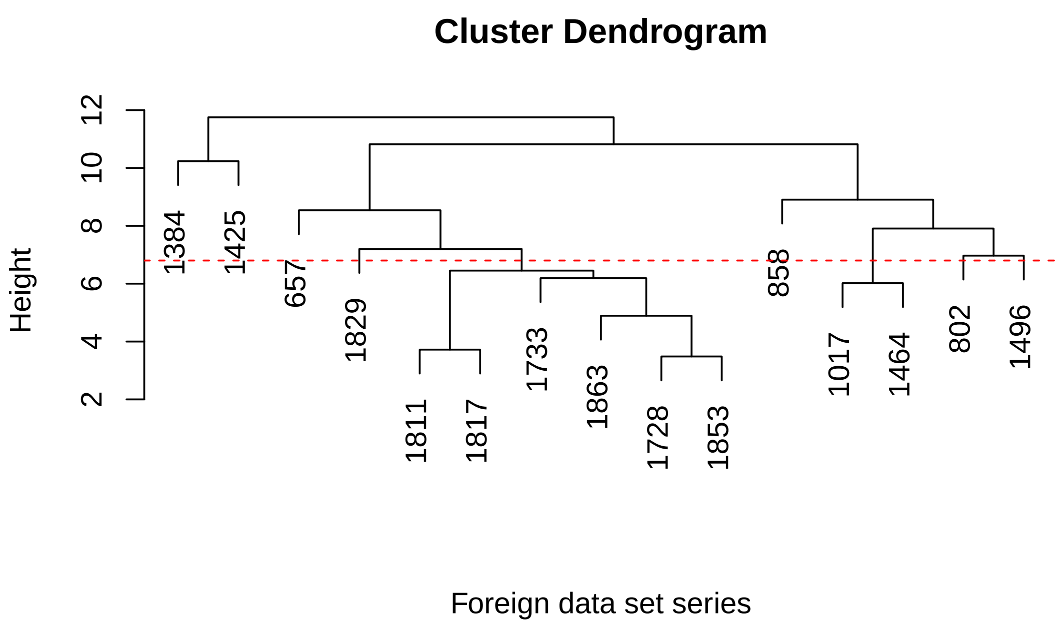

Figure 2.3 *Cluster dendrogram of foreign time series selected after hard thresholding. The labels note the time series code (number)*

The generated dendrogram in figure 2.3 reveals a cluster of 6 time series which happened to be the same series on labour productivity mentioned earlier. The second largest cluster of indicators at the selected level, as marked by a red dashed line, is of size 2 and is formed of series with codes 1017 and 1464. In order to diminish the large cluster to the size of 2 time series, the least angle regression algorithm by Efron *et al* [139] was applied using future growth of coincident economic index as the variable to be forecast. The indicators were ranked according to their predictive power. Next, the least informative indicators were discarded so that the largest cluster diminishes to the size of second-largest cluster.

The results revealed that time series with codes 1817 and 1853 are most informative. They correspond to real labour productivity per person employed in Latvia and in Finland. The other 4 time series from the large cluster were removed and the resulting data set was used in further steps of modelling.

The finalised leading indicators data set was composed of a domestic block which consisted of 4 time series and the foreign block which was formed from 11 series. The number of series constituting the foreign data block is larger in spite of much bigger initial data pool.

The selected indicator set (the full list is given in table 2.2) includes Lithuania's profitable share of enterprises, which was the leading indicator from the domestic leading model [136] which reflects the dynamics in customer purchasing power, labour productivity and the efficiency in management. Foreign direct investment to Lithuania is among selected indicators mostly due to direct causal relationship between investment and future growth of economy; livestock and poultry represent the potential output in the agricultural sector, therefore its presence among selected indicators reveals the importance of agriculture to Lithuanian economy. Lithuania's investment abroad does not have the direct effect on the growth of the economy but it might be a good proxy indicator for business confidence and interest rates[6]. The foreign block included several indicators of con-

[6]Both of these indicators were not considered due to insufficient observations

sumer and business confidence and a few indicators of labour productivity from European countries, a couple of indicators of GDP components from Portugal, Japan and France. The rest of selected leading indicators are net saving of US and gross saving of Cyprus. These indicators reflect fluctuations in financial market: US was selected with regard to its size and enormous impact on international financial sector while Cyprus was selected due to its large offshore banking industry (relative to GDP) and sensitivity to shocks in the finance sector. These results suggest that it might be useful to consider including more financial indicators to initial data set. However, since the financial indicators for Lithuanian economy are few, especially the ones starting at least in 1998, the financial indicators were not included into initial data pool because all financial data would be represented only in foreign block and it could affect the final results by attributing more weight to foreign indicators.

The source of domestic series is Statistics Lithuania, of foreign series - Eurostat.

2.2.2. Dynamic hierarchical factor model

After selecting the leading indicators follows the stage of building a model to evaluate domestic and foreign factors. The method for evaluating the factors is a three level dynamic hierarchical factor model. This method allows to impose a certain structure and estimate separate factors for domestic and foreign variables. The equations constituting the three level hierarchical model are the following (one equation for each hierarchy level) [140]:

$$X_{bit} = \Lambda_{G,bi} G_{bt} + e_{Xbit} \tag{2.15}$$

$$G_{bt} = \Lambda_{F,b} F_t + e_{Gbt} \tag{2.16}$$

$$F_t = \psi F_{t-1} + \varepsilon_{Ft} \tag{2.17}$$

X_{bit} are leading series, which were transformed to be stationary and

Table 2.2 *The final list of selected leading indicators*

Block	Country	Variable
Domestic	Lithuania	Lithuania's investment abroad
Domestic	Lithuania	Foreign direct investment to Lithuania
Domestic	Lithuania	Livestock and poultry
Domestic	Lithuania	A profitable share out of the total number of enterprises
Foreign	Japan	Final consumption expenditure of general government
Foreign	Portugal	Final consumption expenditure of households, total
Foreign	Japan	Household and NPISH final consumption expenditure
Foreign	France	Real Gross Domestic Product per capita
Foreign	Cyprus	Gross saving
Foreign	United States	Net saving
Foreign	Finland	Real labour productivity per person employed
Foreign	Latvia	Real labour productivity per person employed
Foreign	Denmark	Consumer Confidence Index
Foreign	France	Business Confidence Index
Foreign	France	Consumer Confidence Index

scaled (with zero mean and unit variance), index b denotes the block (either domestic or foreign), i - index of time series, t denotes time index. Λ_G and Λ_F are loadings, G_{bt} are block-level factors, F_t is a common factor. The equation (2.17) describes stationary AR(1) process[7]. e_{Xbit}, e_{Gbt} and ε_{Ft} have zero mean and their variances are $\Sigma_X = cov(e_{Xbit})$ and $\Sigma_G = cov(e_{Gbt})$.

Bayesian approach was used because the likelihood function is of complicated form and it might not yield consistent estimation via maximum likelihood method. The evaluation of this model was carried out following the procedure by Moench et al. [15], via Markov Chain Monte Carlo (MCMC) using Gibbs sampling technique (Carter & Kohn [141]), under the assumption of Gaussian innovations.

Data series are structured into blocks $b = 1, 2$, the first one being domestic and the second – supranational. Each series i in a given block b is decomposed into an idiosyncratic component e_{Xbit} and a common component $\Lambda_{G.bi}(L)G_{bt}$ which it shares with other variables in the same block. Each block level factor G_{bjt} has a serially correlated block-specific component e_{Gbjt} and a common component $\Lambda_{F.bj}(L)F_t$ which it shares with all other blocks. The common economy-wide factor F_t is assumed to be serially correlated and follow AR(1) process.

In this model, variables within a block can be correlated through F_t and the e_{Gbjt}'s, but variables between blocks can be correlated only through F_t.

Estimation procedure by MCMC:

Let $\mathbf{\Lambda} = (\Lambda_G, \Lambda_F)$, $\mathbf{\Sigma} = (\Sigma_F, \Sigma_G, \Sigma_X)$.

1. Organize data into blocks to yield $X_{bt}, b = 1, 2$. Use principal components to initialize $\{G_t\}$ and $\{F_t\}$. Use these to produce initial values for $\mathbf{\Lambda}$, ψ and $\mathbf{\Sigma}$.

2. Conditional on $\mathbf{\Lambda}$, ψ, $\mathbf{\Sigma}$ and $\{F_t\}$ draw $\{G_t\}$ taking into account time

[7] The higher order AR(p) processes were considered but modelling showed that coefficients for lags 2 and greater were statistically insignificant

varying intercepts.

3. Conditional on $\mathbf{\Lambda}$, ψ, $\mathbf{\Sigma}$ and $\{G_t\}$ draw $\{F_t\}$.

4. Conditional on $\{G_t\}$ and $\{F_t\}$, draw $\mathbf{\Lambda}$, ψ and $\mathbf{\Sigma}$

5. Return to 2.

The step (1) and step (4) are straightforward. Step (3) could be performed by using Gibbs sampling procedure for dynamic linear models by reducing the 3-level factor model to 2-level factor model constituted by equations (2.16) and (2.17), since the $\{G_t\}$s are "known". The only step that requires a modification in standard methods is step (2). The time varying intercept that has to be conditioned on is a term $\Lambda_{G,bi}\Lambda_{F,b}F_t$ which we get from combining equations (2.15) and (2.16). This term captures the part of the dynamics of the block level factor G_{bt} that it shares with other blocks.

The estimations were carried out using *dlm* (Petris [142]) package of statistical software *R*. Steps (2) and (3) were performed by building 2-level dynamic linear models according to the known parameters and the necessary states were acquired by filtering and sampling the built DLMs using Gibbs procedure, which is implemented in package *dlm*.

10000 iterations were made, and first 500 were dropped out as a 'burn-in'. The choice of number 500 was based on graphical inspection of acquired parameter estimates and following the example of Moench and Ng [16]. The first 700 realisations of $\Lambda_{G,bi}$ elements are plotted in the graph 2.4.

The domestic and foreign leading factors were evaluated calculating the expectation from resulting distributions. Another round of simulations was carried out to compare the results. 100000 iterations were made and first 50000 were discarded. The results are almost identical (mean absolute difference in acquired factors was 0.0034, which is very low since the variance of factors is set to 1). The resulting factors are plotted in graph 2.5.

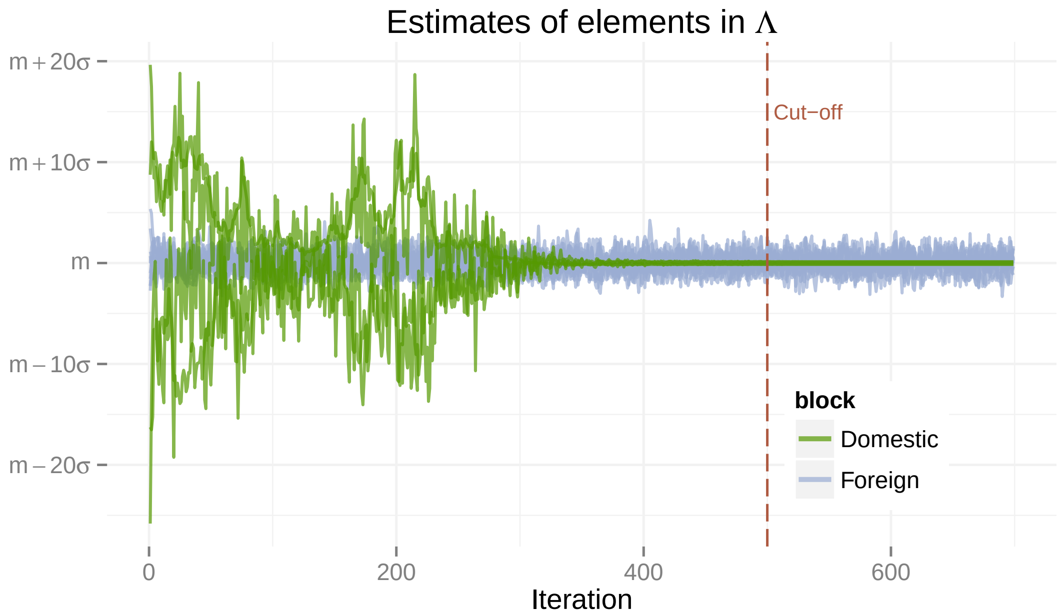

Figure 2.4 *The first 700 realisations of $\Lambda_{G,bi}$ elements. They are scaled to demonstrate the "burn-in", m denotes the mean and σ denotes the standard deviation. It could be identified from this graph that the loadings on the domestic variables converge in about 300–400 iterations and variate very little afterwards.*

The results indicate that even though the extracted domestic and foreign factors are a bit noisy, they depicted the economic crisis and recovery in 2007–2011 pretty well. As expected, domestic and foreign factors have similarities with common factor (the domestic factor $G_{1,t}$ correlates with the common factor by 0.90, foreign factor $G_{2,t}$ correlation with common factor F_t is 0.67). Even though correlation of $G_{1,t}$ and $G_{2,t}$ is positive (0.41) they have periods where they act opposite of each other, which is imminent since the model specification allows them to correlate only through the common factor F_t.

It was assumed that error terms follow normal distribution and initial distributions of $\Lambda_{G,bi}$ elements and ψ values were derived using semi-parametric bootstrap method [143, 144]: repeating the first draw (Gibbs

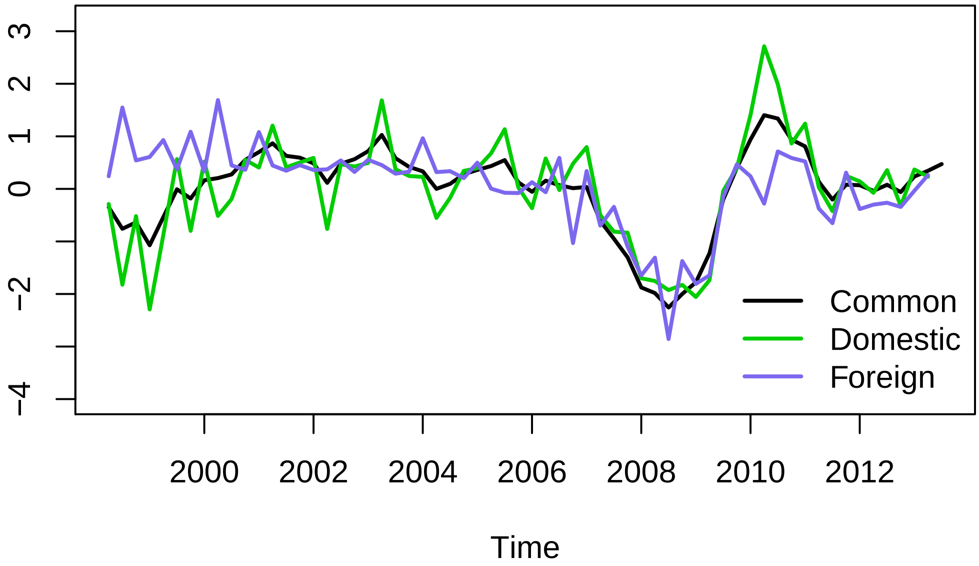

Figure 2.5 *Evaluated common, domestic and foreign leading factors from the hierarchical factor model*

sampling method) 20000 times and calculating the density using gaussian kernel and a 4-times wider bandwidth than the one given by Sheather and Jones' [145] method in order to acquire a smoother shape.

The initial and resulting density estimates of parameter ψ from the equation (2.17) are plotted in graph 2.6. The initial and final density estimates of elements in $\Lambda_{G,bi}$ are plotted in graph 2.7. This graph reveals that 2 elements in $\Lambda_{G,bi}$ have resulting densities, that are centered around zero. Those loadings correspond to domestic variable *Lithuania's investment abroad* and foreign variable *Gross saving in Cyprus*. The results indicate that those variables do not carry relevant information and their loadings are statistically and economically insignificant. They do not have effect on evaluated factors since their average effect is null and resulting mean is calculated from a very large sample. In order to make sure that their effect is insignificant, another round of simulations was carried out excluding these indicators, and resulting factors are nearly identical: mean absolute difference in common factors F_t is 0.006, in domestic factors $G_{1,t}$ is 0.00012 and in foreign factors $G_{2,t}$ is 0.005.

The factors with their filtered confidence intervals are plotted in graph 2.8.

The domestic and foreign blocks of initial and final values of elements in covariance matrix Σ_X are plotted in graphs 2.9 and 2.10.

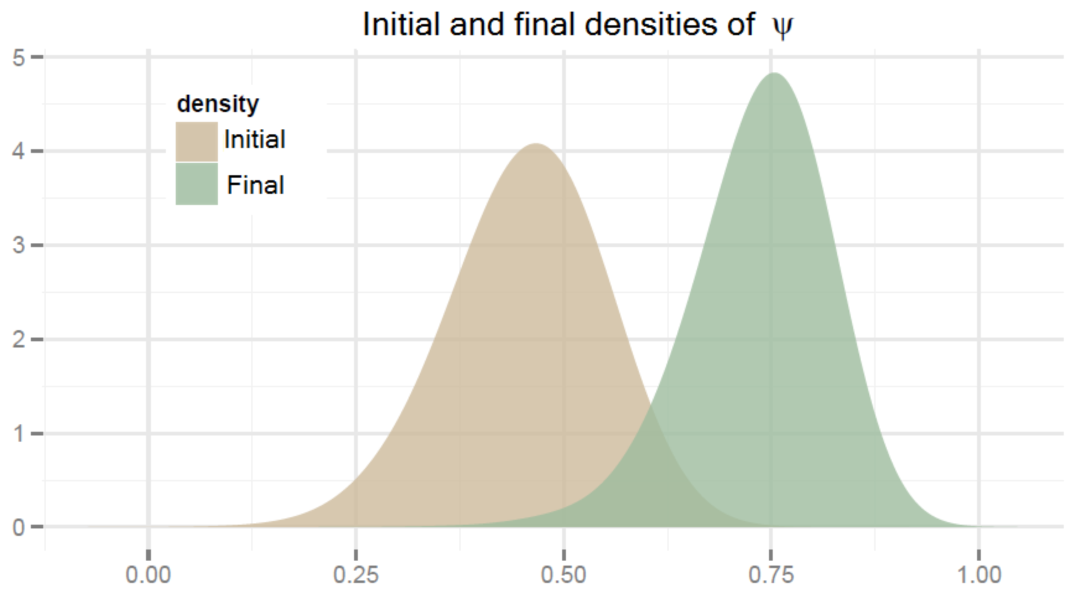

Figure 2.6 *The initial and final densities of the parameter ψ (the autoregressive coefficient in a third-level equation* (2.17)*)*

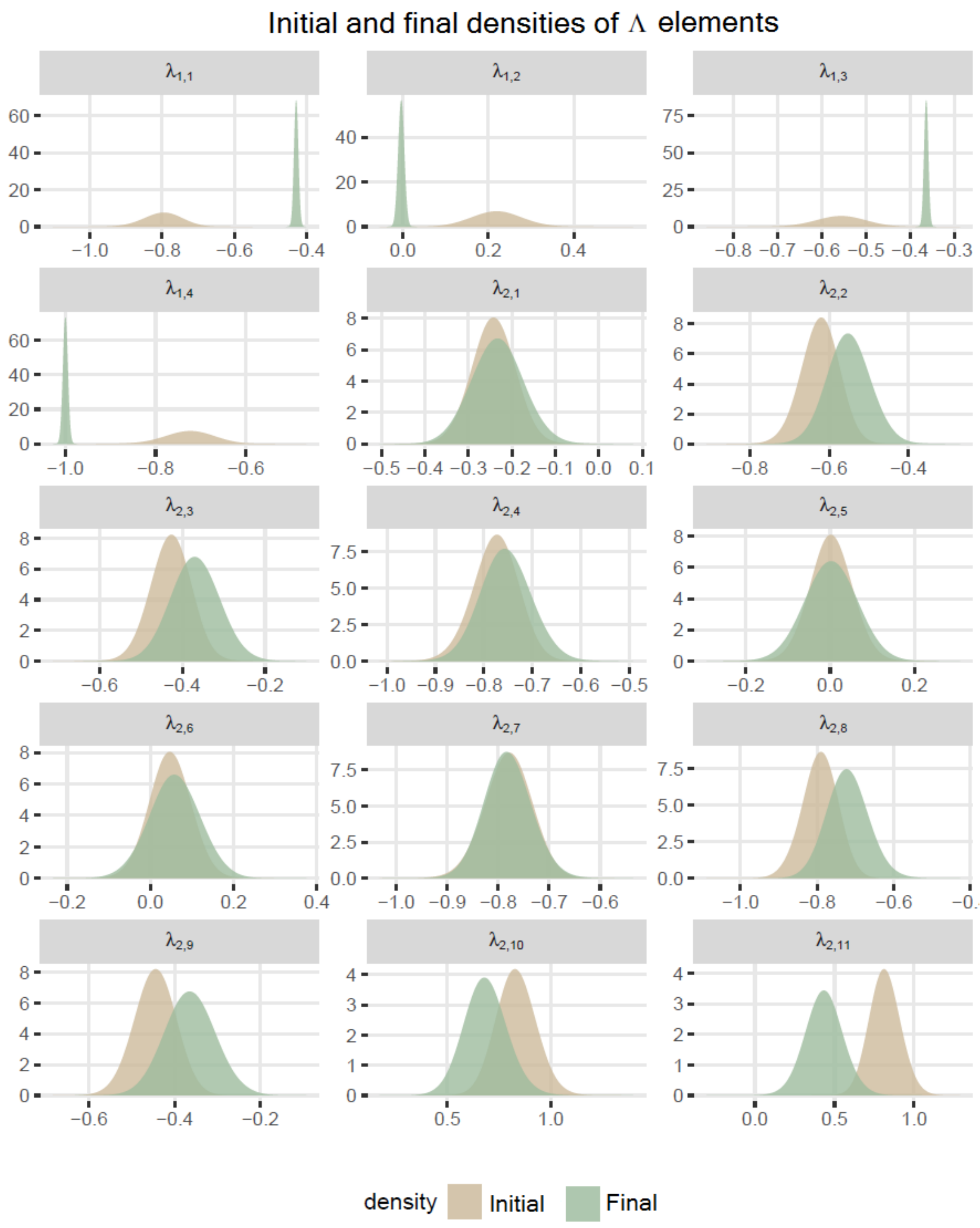

Figure 2.7 *The initial and final densities of the elements in* $\Lambda_{G,bi}$ *(loading vectors in first-level equation* (2.15)*). This graph reveals that the final densities of loadings on domestic indicators have very small variance.*

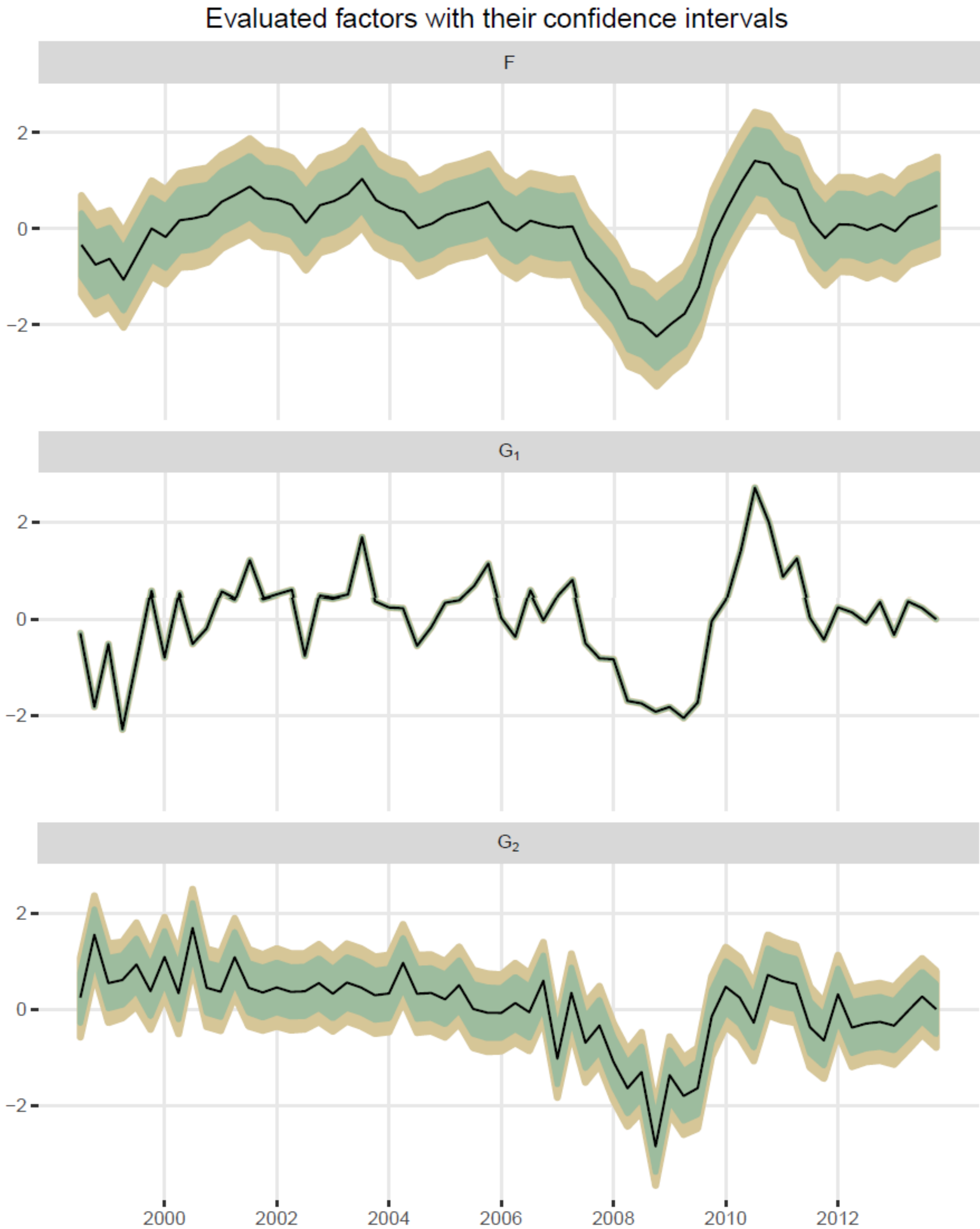

Figure 2.8 *Evaluated common, domestic and foreign leading factors with their resulting confidence intervals of level 0.8 and 0.95. It could be identified from this graph that domestic factor G_1 variates very little since the confidence intervals for filtered states are very narrow.*

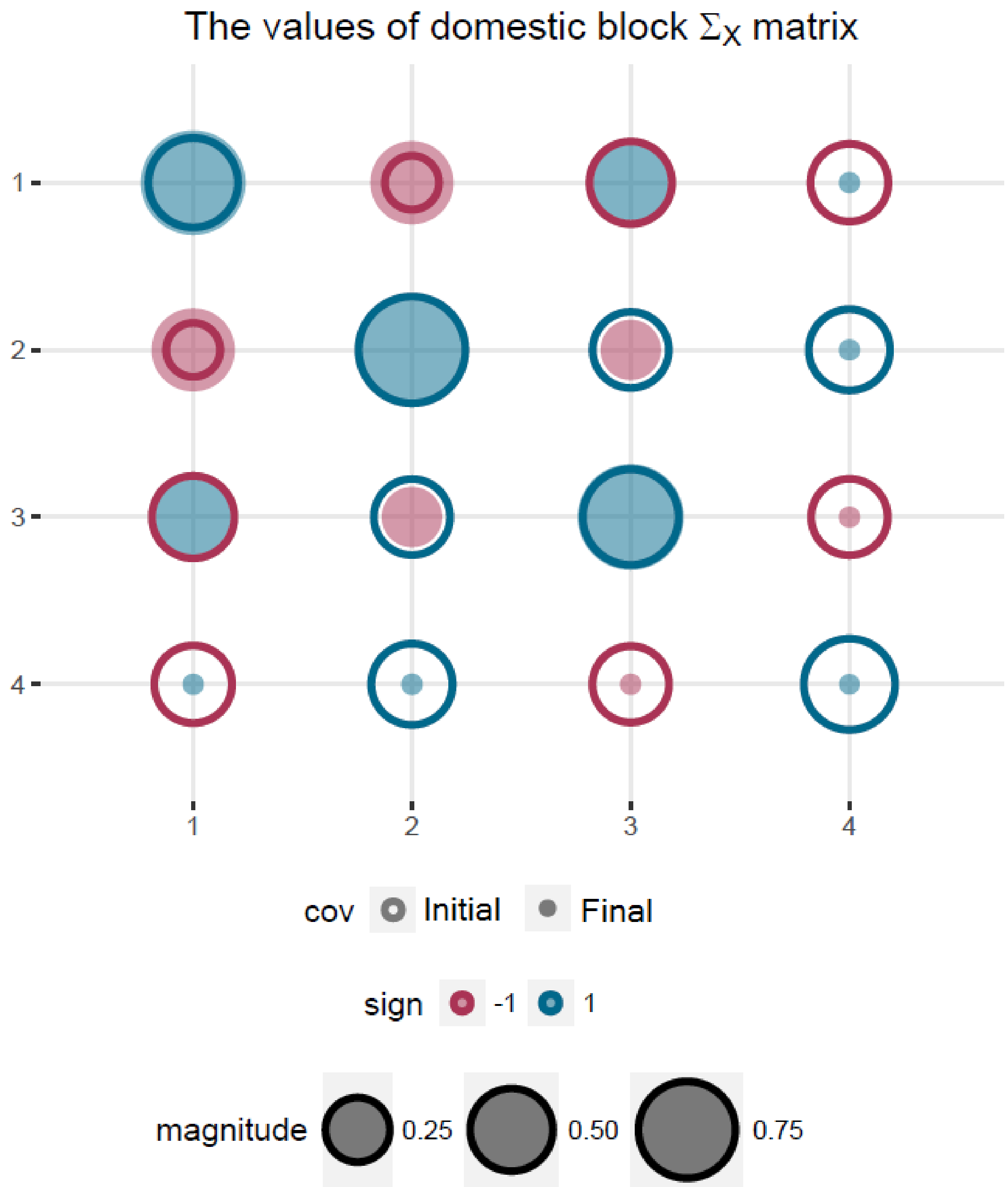

Figure 2.9 *The elements of initial and final covariance matrices Σ_X: the domestic block. The covariance associated with the fourth variable of domestic block (the proportion of profitable enterprises) shrunk significantly implying that this indicator was the most informative in this block.*

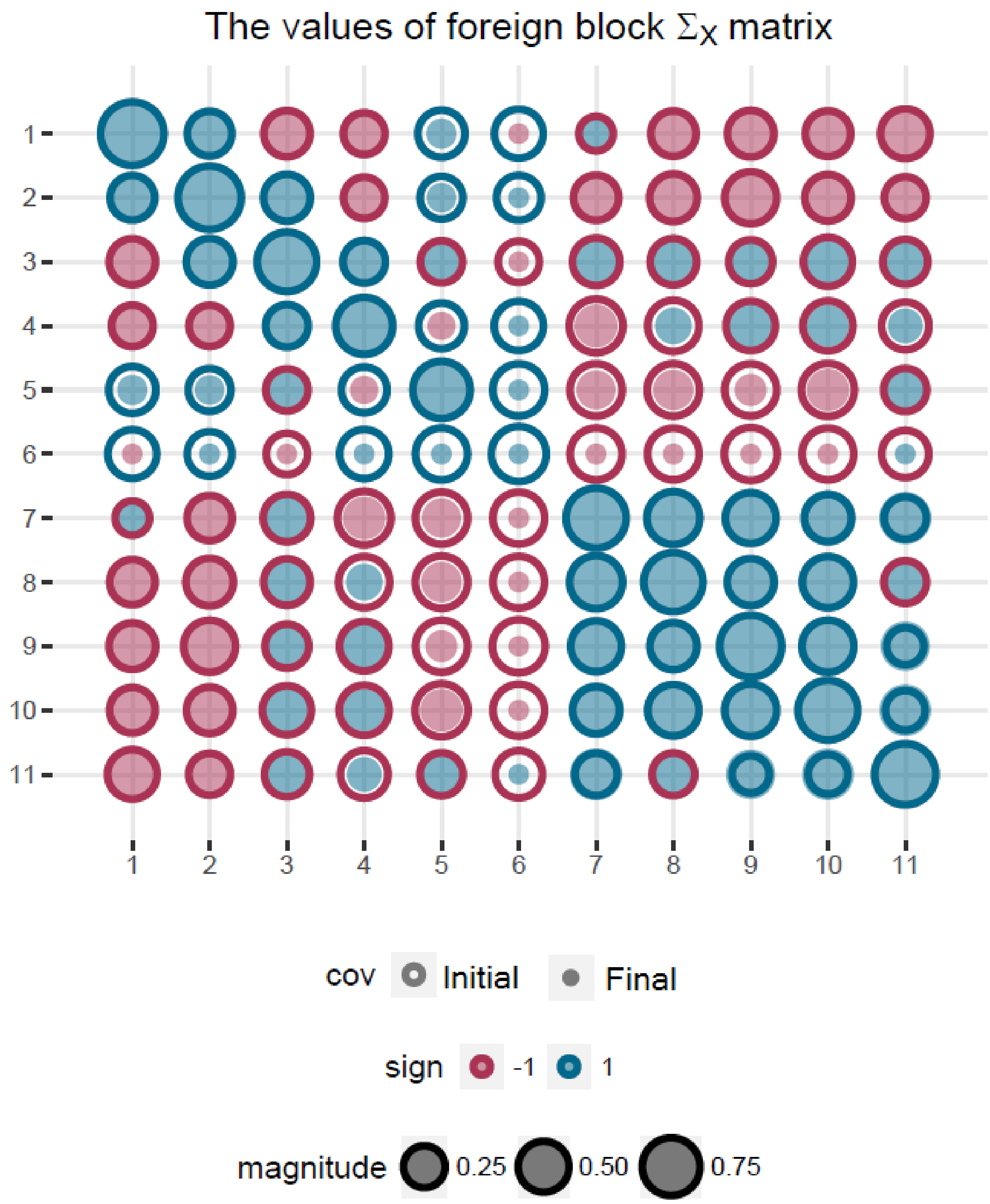

Figure 2.10 *The elements of initial and final covariance matrices Σ_X: the foreign block. The covariance associated with the sixth variable of domestic block (the GDP per capita of France) shrunk significantly implying that this indicator was the most informative in this block.*

2.2.3. Varying factor load evaluation

The proposed measure of globalisation relies on evaluating the portion of economic growth explained by foreign and international indicators in comparison to domestic ones. In order to capture the load of domestic and foreign indicators on the future growth of Lithuanian economy a linear model following the idea of Stock and Watson [131] was considered in the form of regressing the growth of coincident index on both leading factor estimates. Let us define $Y_t = \frac{C_{t+2}}{C_t}$ and scale it to have zero mean and unit variance. Afterwards, the regression (2.18) is evaluated:

$$Y_t = \alpha_1 G_{1,t} + \alpha_2 G_{2,t} + \varepsilon_t \tag{2.18}$$

Here $G_{1,t}$ is a the domestic leading factor, and $G_{2,t}$ is the foreign leading factor. The expression from equation (2.12) was reduced to (2.18) based on statistical significance of parameters in linear regression. The estimates of equation (2.18) parameters are given in table 2.3.

Table 2.3 *Estimates of model parameters from equation* (2.18) *describing the average load of domestic* (α_1) *and foreign* α_2 *variables on the future growth of economy represented by coincident economic index*

	Estimate	Std. Error	t value	p-value
α_1	0.351	0.117	2.994	0.004
α_2	0.281	0.117	2.395	0.020

Precision One may raise the question if factors from the hierarchical model provide better results in terms of accuracy than previous attempt by Reklaite [136] based on including leading indicators directly into the forecast equation. In 2011 paper adjusted R^2 for CEI growth predicion of the regression was 0.524. The update on the same regression (using larger data span) gives R^2 of value 0.401. The determination coefficient (adjusted R^2) of new regression (2.18) is 0.398. However cross-validation (1 step ahead prediction was built for time period t from 2001 to 2013 fitting model using observations up to period $t-1$) shows better precision by the new method:

Measure	ME	RMSE	MAE	MPE	MAPE
Reklaite (2011)	-1.418	3.802	2.627	741.15	788.96
New method	0.2025	3.397	2.541	538.76	546.07

Since we are more interested in the dynamics of α parameters, the equation (2.18) has to be modified to include time-varying coefficient on factors.

Assumptions

Let us assume that the part of economic growth forecast explained by foreign and domestic indicators is time-invariant, i.e.

$$\alpha_{1,t} + \alpha_{2,t} = \gamma$$

Another assumption that we are going to make is that the "true" parameters $\alpha_{1,t}$ and $\alpha_{2,t}$ are varying in time, so that the variance is composed of time-independent element and time-depending component. Therefore, in the linear regression (2.18) the variance of estimates of α_1 and α_2 could also be de-composed:

$$\mathrm{Var}(\hat{\alpha_b}) = \mathrm{bias}_t(\alpha_b) + \mathrm{Var}(u_t)$$

here $b = 1, 2$, $u_t \sim i.i.d.$

The intent is to model the time-depending bias as autoregressive process. Therefore a dynamic linear model was considered:

$$Y_t = \alpha_{1,t} G_{1,t} + (\gamma - \alpha_{1,t}) G_{2,t} + \varepsilon_t, \tag{2.19}$$

$$(\alpha_{1,(t+1)} - c) = \phi(\alpha_{1,t} - c) + u_t. \tag{2.20}$$

c is the mean of AR(1) process from (2.20). From (2.20) we can derive the variance of $\alpha_{1,t}$:

$$\mathrm{Var}(\alpha_{1,(t+1)}) = \phi^2 \mathrm{Var}(\alpha_{1,t}) + \mathrm{Var}(u_t)$$

$$\mathrm{Var}(\alpha_{1,(t+1)}) = \frac{\mathrm{Var}(u_t)}{1-\phi^2}$$

Using (2.18) we have estimate $\widehat{\mathrm{Var}}(\hat{\alpha_i}) \approx 0.0137$ which will be used as a restriction for parameters ϕ and $\mathrm{Var}(u_t)$:

$$0.0137(1-\phi^2) = \mathrm{Var}(u_t)$$

Another assumption is setting $\mathrm{Var}(\varepsilon_t)$ from equation (2.19) to match the error variance from (2.18) regression estimates.

Under the model specification with equations (2.19) and (2.20) the mean of AR(1) process denoted by c should match the parameter α_1 estimate from equation (2.18).

The last assumption that is needed to make is about distribution of $\alpha_{1,t}$ at the starting time point $t = 0$. For that reason a series of regressions from subsets of data are run using a moving-window approach. From (2.19) we get

$$Y_t - \gamma G_{2,t} = \alpha_{1,t}(G_{2,t} - G_{2,t}) + \varepsilon_t \tag{2.21}$$

The evaluation was carried out using local linear regression method with uniform kernel on a window covering 5 year period (20 observations). Regressions were run with quarterly shift of 5-year span window starting with 1998, i.e. first window covered time span from 1998 to 2003, the last window covered time span from 2008 to 2013. The resulting $\alpha_{1,t}$ estimates with their 80% confidence intervals are plotted in graph 2.11. Even though the standard errors are quite large, the it could be detected that there has been a shift in parameter $\alpha_{1,t}$ and the domestic impact on economic growth has been declining. The estimates from the window covering the earliest period give $\hat{\alpha_1} \approx 0.584$ and the standard error $\widehat{\text{se}}(\hat{\alpha_1}) \approx 0.133$. These values will be used as a prior information on $\alpha_{1,t}$ at the starting time point $t = 0$ in the dynamic linear model estimation.

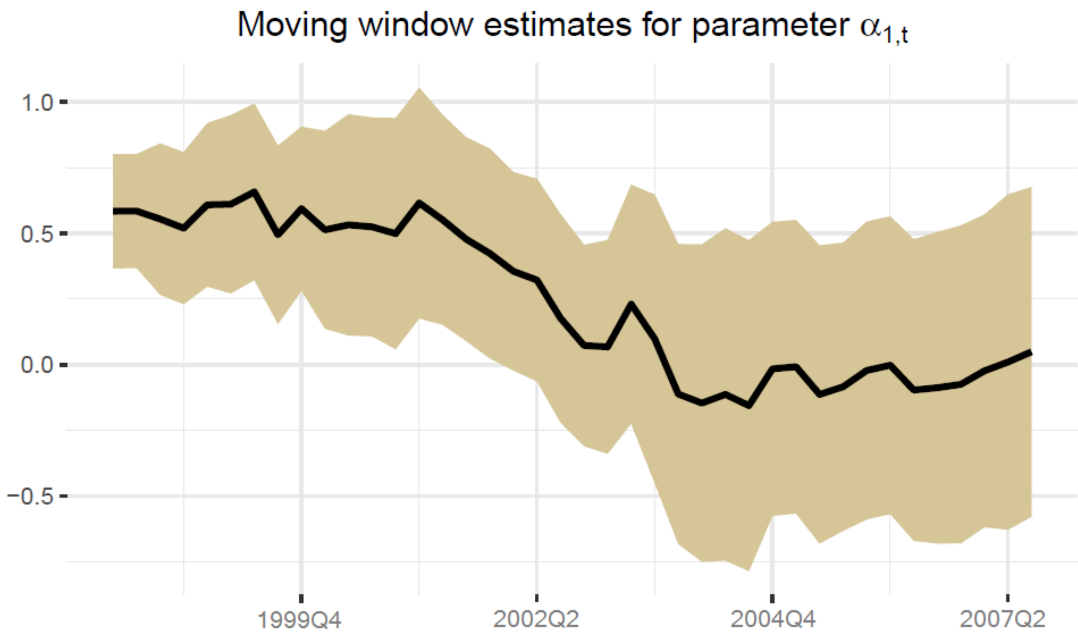

Figure 2.11 *Local linear regression estimates of $\alpha_{1,t}$ - the load of domestic factor impact on the future growth of economy and 0.8 level confidence band*

Dynamic linear model

The parameters of the dynamic linear model given by equations (2.19) and (2.20) were evaluated using previously described assumpions and applying maximum likelihood method assuming that innovations ϵ_t and u_t are gaussian.

This model was built assuming that Y_t, $G_{1,t}$ and $G_{2,t}$ are given, i.e. observed series and that $\alpha_{1,t}$ is a state which has to be filtered to get its estimate. Maximum likelihood estimate for equation (2.20) parameters: $\hat{\phi} \approx 0.89$ and $\widehat{\text{Var}}(u_t) \approx 0.0027$. Filtered $\alpha_{1,t}$ series with its 90% confidence band is given in Fig. 2.13.

The hypothesis that we are trying to validate is that the proportion

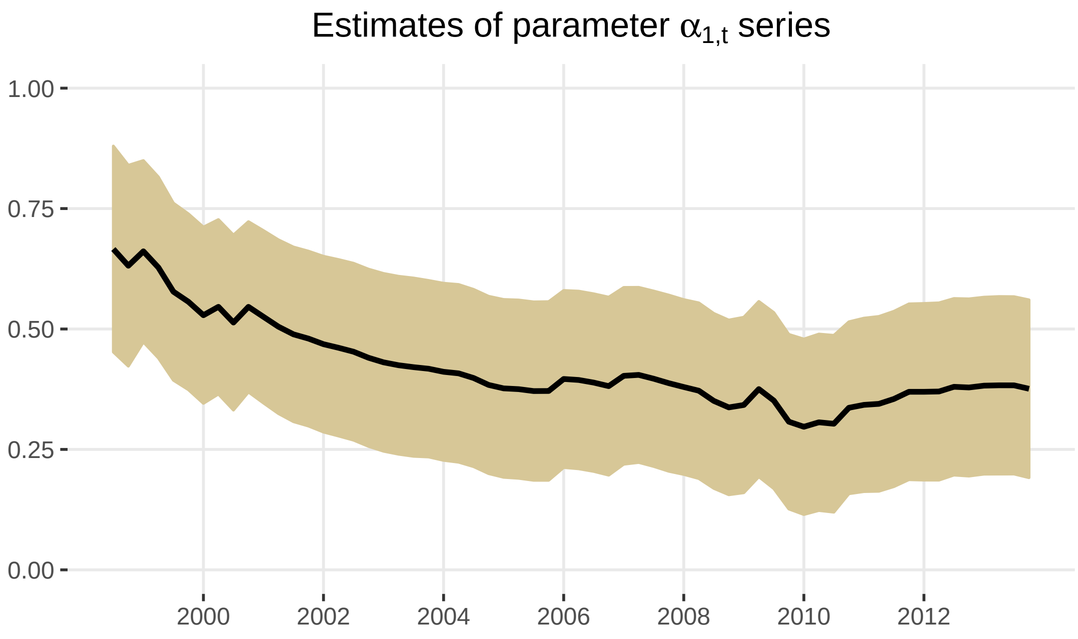

Figure 2.12 *Evaluated parameter series $\alpha_{1,t}$ - the load of domestic factor impact on the future growth of economy and 0.9 level confidence band*

of economic growth forecast explained by foreign indicators is increasing over time. Under this specification this hypothetical statement means that parameter $\alpha_{1,t}$ should be decreasing over time. The results from the graph 2.12 show that the importance of domestic variables indeed diminished over time.

Globalisation measure

Proposed definition of a new globalisation measure describes it as a portion of the future economic growth explained by foreign indicators in comparison to domestic ones. The results of our model imply that this measure is $\gamma - \alpha_{1,t} = \alpha_{2,t}$. The estimate of proposed globalisation measure is given is Fig. 2.13.

It can be identified from the graph 2.13 that globalisation measure estimate is increasing, which means that Lithuanian economy is more and more

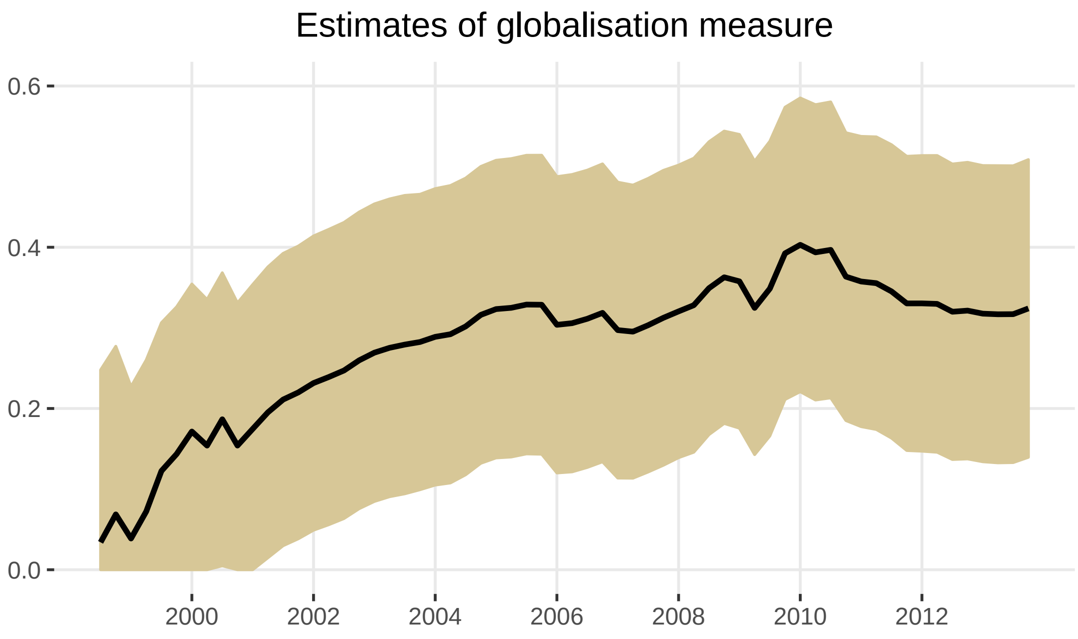

Figure 2.13 *Evaluated parameter series of globalisation measure - the load of foreign factor impact on the future growth of economy and 0.9 level confidence band*

intertwined with foreign economies. This result also validates the hypothesis about the increasing amount of forecast explained by foreign indicators. It leads to a conclusion that globalisation can be measured by the proposed indicator and its effect on focal economy is increasing in magnitude over time.

Comparison with results of other authors

The comparison of these results to findings of other researchers is limited since majority of studies focus on a single time period and ranking of countries, and the rest rarely include Lithuania. The globalisation measures for Lithuania that cover more than one time point:

Globalisation measure	2000	2001	2002	2003	2004	2005	2006
KOF index [80]	–	–	–	67.55	69.76	70.19	71.07
Maastricht index [82]	43.99	–	–	–	–	–	–
CSGR index [79]	0.147	0.158	0.194	0.210	0.253	–	–

Globalisation measure	2007	2008	2009	2010	2011	2012	2013
KOF index	72.93	72.45	68.80	72.06	73.22	72.83	77.26
Maastricht index	–	59.89	–	–	–	61.74	–

The table indicates that Lithuania tends to increase its international integration over time and these results are consistent with $\alpha_{2,t}$ series estimate. $\alpha_{2,t}$ estimate is denser in time domain, since it is a quarterly measure, therefore it provides the possibility to inspect short-term developments as well as long-term trend.

2.2.4. Structure validation

The acquired domestic and foreign factors have a desired interpretation but one might want to validate if the imposed structure is statistically justified. In order to validate the imposed structure another factor model was built which had 2 factors in a single block, i.e. domestic and foreign leading series were pooled together and 2 dynamic factors were evaluated from that data set. The evaluation followed the same method as 3-level factor model only the algorithm used had one step less because the single-block factor model had only 2 levels. The dynamics of evaluated factors were assumed to follow AR(1) process, the same order as the common factor in 3-level factor model. Single-block factors are orthogonal of each other — it is required by model specification.

The (empirical) correlation matrix of acquired factors from structural approach $G_{1,t}$, $G_{2,t}$ and factors from non-structural approach $F_{1,t}$ and $F_{2,t}$ is in the table 2.4.

It can be identified that even without the imposed block structure the

Table 2.4 *Correlations between factors acquired from structural ($G_{1,t}$ and $G_{2,t}$) and non-structural ($F_{1,t}$ and $F_{2,t}$) approach*

	$G_{1,t}$	$G_{2,t}$	$F_{1,t}$	$F_{2,t}$
$G_{1,t}$	1	0.41	0.05	0.96
$G_{2,t}$	0.41	1	0.72	0.43
$F_{1,t}$	0.05	0.72	1	0.05
$F_{2,t}$	0.96	0.43	0.05	1

factors from structural approach correlate with 2 factors from non-structural approach by 0.96 and 0.72. This means that the information of series from 2 different blocks naturally form 2 different factors. The structural approach lets us name those factors and give them interpretation which could be very difficult to justify in the case of non-structural factors.

In order to measure the statistical fit, Akaike ($AIC(M) = \log(\sigma^2(M)) + 2\frac{k}{N}$) and Schwarz ($BIC(M) = \log(\sigma^2(M)) + k\frac{\log(N)}{N}$) information criteria were estimated for a non structural 2-factor model (2FM), and the dynamic hierarchical factor model (DHFM). Here M denotes the model, k is a number of parameters, $N = m \times T$ is a number of data points, T — a number of time points (quarters), and m is a number of indicators.

	2FM	DHFM
AIC	-0.364	-0.344
BIC	-0.178	-0.239

Akaike criterion slightly favours the non-structural model, but Schwartz criterion indicates the hierarchical approach as more precise. The regression (2.18) with 2FM factors gives adjusted $R^2 = 0.223$ which is considarebly inferior to the adjusted $R^2 = 0.398$ given by DHFM factors.

Table 2.5 *Estimates of model parameters from equation* (2.18) *describing the average load of 2 factors acquired from non-structural approach, on the future growth of economy represented by the coincident economic index*

	Estimate	Std. Error	t value	p-value
α_1	0.414	0.112	3.686	0.000498
α_2	0.275	0.126	2.191	0.03239

These characteristics indicate that the structural approach is better justified statistically and gives a greater precision.

2.2.5. Remarks on the results

Globalisation indexes are vulnerable to the critique that they measure international integration without distinguishing regionalisation from globalisation. Only trade-based measures deal with this issue since trade data could be weighed on geographical distance. Other methods do not provide this option since syndicate measures require the data that does not have division into countries (e.g. internet bandwidth or international calls) or it is very sparse. The proposed method could deal with this issue if foreign block is divided into sub-blocks using geographical division, e.g. European countries vs. non-European countries. However, in case of Lithuanian indicators, the foreign block is formed from 11 indicators, and only 3 of them are non-European indicators representing 2 countries. Upon attempt to apply the division the results revealed that data is not sufficient to make inference from. The 0.9 level confidence interval on non-European load on foreign factor is $(0.087, 0.257)$ and European load is $(-1.214, -0.936)$. One would expect them to be of the same sign, but it is the oposite. This could be explained by different nature of indicators and low amount of information. Therefore the non-European block has insufficient data to distinguish Lithuanian international integration into regionalisation and globalisation. Nevertheless, the method is available and could be applied on extended data set including more indicators from non-European countries.

The resulting densities of elements in Λ_G which are plotted in graph 2.7 are consistent with economic rationale to a large extent. The most informative parameter from domestic block is a profitable share out of the total number of enterprises, which confirms the previous findings [136]. The loadings in the foreign block are of the same sign except for net saving in U.S., consumer and business confidence indices in France. The saving in U.S. has adverse relation to economic growth, but loading coefficients on France confidence indices are unexpected. Upon additional inspection it was revealed that Denmark and France use a different methodology to estimate them. Moreover, indicators in France have a high volatility and short-term fluctuations, which might have caused resulting parameter estimates. These results suggest that it might be a good idea to use an additional leading indicator selection criterion and inspect if the parameter acquired regressing the coincident economic index on a potential leading indicator has an economically justified sign. A re-run of simulations excluding confidence indices of France shows that factors changed little — mean absolute difference in foreign factors is 0.072, in domestic ones is 0.04 and in common factor is 0.136.

2.3. Economic growth prediction on GDP

Having developed the globalisation measure which is based on apportioning the part of growth explained by foreign and international indicators on future growth of economy which is represented by CEI, one might want to examine if similar results would be acquired by using the growth of GDP as a measure of economic activity. Schumacher [4] concluded that including targeted international predictors shows promising results in forecasting the GDP for Germany. If the foreign effect could be observed for an economy so large as Germany, the effect of foreign variables should be even more evident modelling it on Lithuania.

The leading variable selection was performed on the same initial data set as in section 2.2. The only difference was that the growth of GDP was a variable of interest in the selection criteria:

1. Granger causality between potential leading series and growth of GDP (pairwise testing for lag depth 2 with significance level $\alpha = 0.05$)
2. Correlation between potential leading series $\Delta X_{i,(t-l)}$ and real GDP growth ΔGDP_t should be greater with lags $l > 0$

The list of selected leading indicators is very similar to the list in table 2.2 and is given in table 2.6

The dynamic hierarchical factor model was built using the equations (2.15), (2.16), (2.17) and assumptions presented in section 2.2.2. The evaluation was performed following the same algorithm and the evaluated factors are plotted in figure 2.14. The resulting factors are very similar to the ones in figure 2.5, especially the domestic factor since the domestic block consisted of the same indicators for both cases. As expected, domestic and foreign factors have similarities with common factor (domestic factor $G_{1,t}$ correlates with common factor by 0.88, foreign factor $G_{2,t}$ correlates with common factor F_t by 0.56). Correlation between $G_{1,t}$ and $G_{2,t}$ is 0.29.

In order to evaluate the average effect of domestic and foreign variables on the future growth of economy, a linear regression was built following Stock and Watson [131]:

$$\Delta GDP_{t+1} = \alpha_1 G_{1,t} + \alpha_2 G_{2,t} + \varepsilon_{t+1} \tag{2.22}$$

Estimates of model parameters from equation (2.22) are given in table 2.7:

Table 2.6 *The selected leading indicators using selection criteria with real growth of GDP as a variable of interest*

Block	Country	Variable
Domestic	Lithuania	Lithuania's investment abroad
Domestic	Lithuania	Foreign direct investment to Lithuania
Domestic	Lithuania	Livestock and poultry
Domestic	Lithuania	A profitable share out of the total number of enterprises
Foreign	OECD - Total	Private final consumption expenditure
Foreign	Japan	Final consumption expenditure of general government
Foreign	Portugal	Final consumption expenditure of households, total
Foreign	Japan	Household and NPISH final consumption expenditure
Foreign	France	Real Gross Domestic Product per capita
Foreign	Cyprus	Gross saving
Foreign	United States	Net saving
Foreign	Euro area (17 countries)	Real labour productivity per hour worked
Foreign	European Union (27 countries)	Real labour productivity per hour worked
Foreign	Finland	Real labour productivity per hour worked
Foreign	Latvia	Real labour productivity per hour worked
Foreign	Finland	Real labour productivity per person employed
Foreign	Latvia	Real labour productivity per person employed
Foreign	Euro area (12 countries)	Self-employed - national concept
Foreign	European Union (15 countries)	Self-employed - national concept
Foreign	Estonia	Business Confidence Index
Foreign	Austria	Consumer Confidence Index
Foreign	Denmark	Consumer Confidence Index
Foreign	France	Business Confidence Index
Foreign	France	Consumer Confidence Index

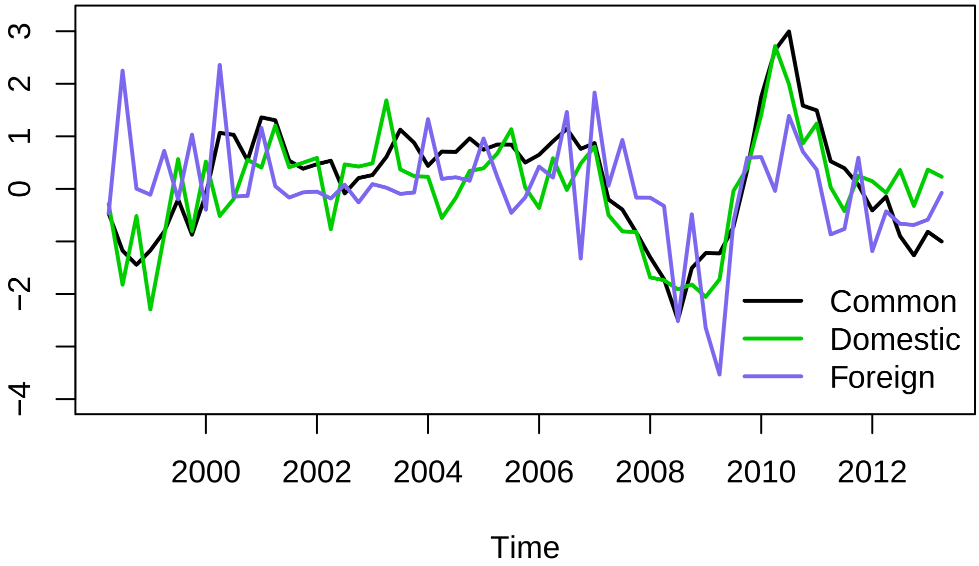

Figure 2.14 *Evaluated foreign, domestic and common factors from selected leading indicators on the GDP as a variable of interest*

The estimates from table 2.7 were used to set the initial value for dynamic linear model evaluation. Since the variable of interest is a portion of forecast on future growth of economy explained by international indicators in comparison to domestic ones, the equations constituting the DLM were introduced the same assumptions building the constraints as for equation (2.19):

$\gamma = 0.696$, $\mathrm{Var}(\varepsilon_t) = 0.689$, $\mathrm{Var}(u_t) = 0.0125 \cdot (1 - \phi^2)$, $\alpha_{1,t=0} \sim \mathcal{N}(0.636, 0.016)$

$$\Delta GDP_{t+1} = \alpha_{1,t} G_{1,t} + (\gamma - \alpha_{1,t}) G_{2,t} + \varepsilon_t, \tag{2.23}$$

$$\alpha_{t+1} = \phi \alpha_{1,t} + u_t. \tag{2.24}$$

Maximum likelihood estimation revealed that $\hat{\phi} = 0.04$ which is statistically insignificant. Therefore a conclusion was made that in this case $\alpha_{1,t}$ is

Table 2.7 *Estimates of parameters in equation* (2.22) *describing average load of domestic* (α_1) *and foreign* (α_2) *variables on the future growth of economy as measured by GDP*

Coefficient	Estimate	St. error	t-value	p-value
α_1	0.297	0.113	2.639	0.010
α_2	0.399	0.113	3.543	0.001

unlikely to follow a stationary process and the DLM was modified to have a random walk in transition equation, i.e. setting $\phi = 1$. Maximum likelihood estimate for $\mathrm{Var}(u_t)$ is 0.0015. Filtered series of a time-varying coefficient from DLM (2.23) were acquired and $\alpha_{2,t} = \gamma - \alpha_{1,t}$ were estimated. The resulting globalisation measure with 0.9 level confidence band is presented in figure 2.16.

The filtered $\alpha_{2,t}$ series do not show a clear trend. However, it can be seen that the effect of foreign variables on the growth of economy has risen over time. Also, it is depicted that foreign indicators had an increasing effect in economic crisis and the recovery in 2008–2011. Since the crisis of 2008–2009 was global, these results agree with our globalisation measurement.

Even though the graph 2.16 shows an increment in foreign effect on the focal economy, especially in the period of global economic crisis, the graph 2.13 shows stronger indication that foreign impact on future growth on economy is increasing. It could be concluded that CEI is more plausible selection when inspecting the globalisation impact on economic growth from theoretical view as it captures multi-domain developments of economic activity and is consistent with the definition of the phenomenom. The empirical research reveals that the globalisation measurement based of foreign effect on GDP forecast is more focused on economic channel of international integration since it showed increased foreign influence in the period of the global economic crisis.

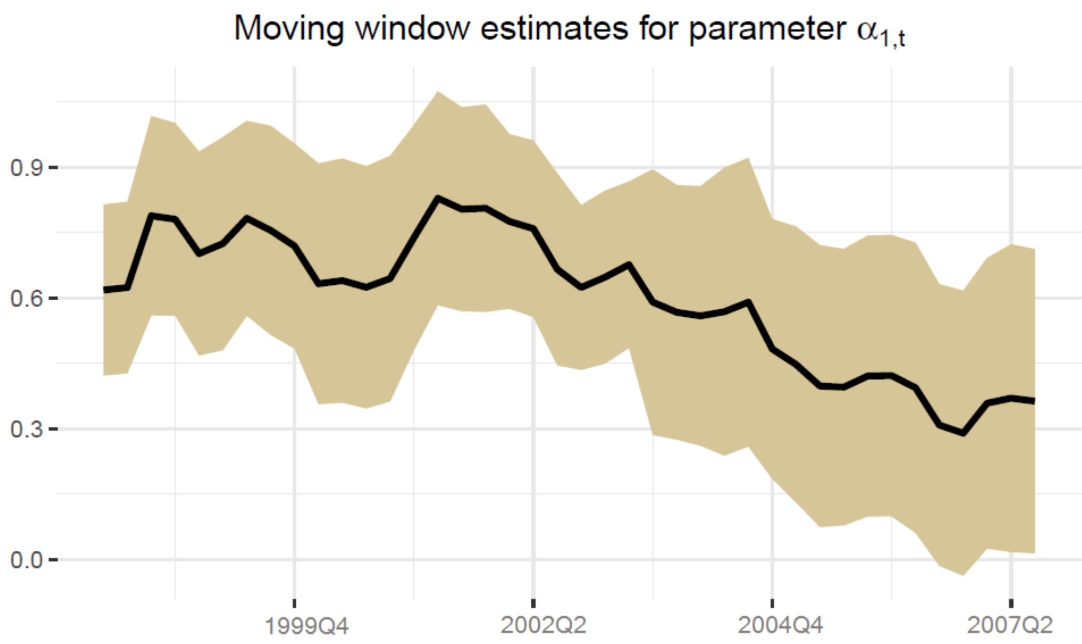

Figure 2.15 *Evaluated moving-window estimates for α_1 from equation (2.22) depicting the portion of GDP explained by domestic indicators 0.9 level confidence band*

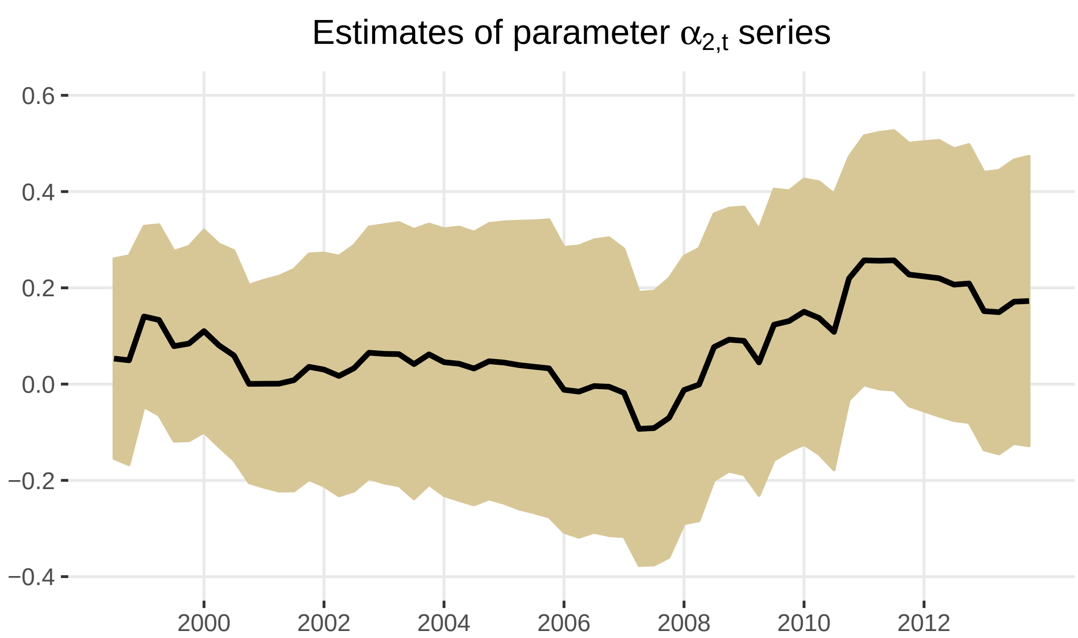

Figure 2.16 *Evaluated α_t series from equation (2.23) depicting the portion of GDP explained by foreign indicators relative to domestic ones and its 0.9 level confidence band*

2.4. Chapter summary

A new globalisation measure was offered. It is based on evaluating the amount of forecast economic growth explained by foreign indicators in comparison to domestic ones. The economic growth is represented by growth of coincident economic index in order to capture the multi-sectoral developments in focal economy.

The methodology for evaluating this measure was developed and the main stages of it are the following:

1. Coincident economic index is evaluated.
2. The selection of leading indicators is performed.
3. Using the selected leading indicators the dynamic hierarchical factor model is built. The domestic and foreign factors are evaluated.
4. Additional analysis is performed in order to set restrictions and assumptions on the parameters of the dynamic linear model.
5. The dynamic linear model is built and the results are used to evaluate a time varying load of foreign indicators on the future growth of economy. The acquired series represent the globalisation measure.

Conclusions

The main conclusions of this thesis are the following:

1. A new globalisation measure is proposed which has a clear interpretation, captures a multi-dimensional nature of globalisation process and is denser in time-domain than the majority of other measures.

2. The methodology has been developed to evaluate the proposed globalisation measure. It relies on fitting the design of a dynamic hierarchical factor model (DHFM) to suit the structure of the data in order to enable the evaluation of effect of grouped indicators on the variable of interest.

3. It was assessed that the factors from DHFM are no less informative than factors acquired from non-structural approach. In addition to that DHFM factors extract more relevant information for the forecast — it was demonstrated using a linear regression with time-invariant coefficients.

4. The empirical research on Lithuanian economy revealed that the parameter on the forecast equation has a trend. This result indicates that ignoring a time-varying nature of parameters might lead to a forecast bias and cause the diminishing prediction accuracy.

5. The results from the dynamic linear model show that the portion of future economic growth explained by foreign indicators relative to domestic ones is increasing for Lithuanian economy and it reflects the globalisation effect.

Discussion

The strong feature of the proposed method is the flexibility in ways to impose the structure and restrictions, therefore various set-ups for interpretation could be built. For example, if more levels of the hierarchy are introduced into the model of domestic and foreign leading indicators, the sub-blocks could be organised using geographical division. This way one could examine the foreign component with more detail and identify the key contributors to the globalisation process.

Using the same approach other problems could be addressed since the method is universal. Any prediction task using a large number of predictors could be used to identify the proportions of the underlying structure of the forecast. As long as the division of data is justified from economic point of view, the analysis should produce sensible interpretation.

Bibliography

[1] K. Drechsel and R. Scheufele. Should we trust in leading indicators? Evidence from the recent recession. Discussion paper 10, Halle Institute for Economic Research, 2010.

[2] A. Reklaitė. Globalisation effect measure via hierarchical dynamic factor modelling. *Equilibrium*, 10(3):139–149, 2015.

[3] F. Fichtner, R. Rüffer, and B. Schnatz. Leading indicators in a globalised world. Working paper 1125, European Central Bank, 2009.

[4] C. Schumacher. Factor forecasting using international targeted predictors: the case of German GDP. *Economics Letters*, 107(2):95–98, 2010.

[5] M. Brisson, B. Campbell, and J.W. Galbraith. Forecasting some low-predictability time series using diffusion indices. *Journal of Forecasting*, 22(6–7):515–531, 2003.

[6] S. Eickmeier and T. Ng. Forecasting national activity using lots of international predictors: An application to New Zealand. *International Journal of Forecasting*, 27(2):496–511, 2011.

[7] A. Banerjee, M. Marcellino, and I. Masten. Leading indicators for euro-area inflation and GDP growth*. *Oxford Bulletin of Economics and Statistics*, 67(s1):785–813, 2005.

[8] S. Eickmeier and C. Ziegler. How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting*, 27(3):237–265, 2008.

[9] James H Stock and Mark W Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179, 2002.

[10] J.H. Stock and M.W. Watson. Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23(6):405–430, 2004.

[11] G. Rünstler, K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, K. Ruth, and C. Van Nieuwenhuyze. Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise. *Journal of Forecasting*, 28(7):595–611, 2009.

[12] E. Angelini, G. Camba-Mendez, D. Giannoneand L. Reichlin, and G. Rünstler. Short-term forecasts of euro area GDP growth. *The Econometrics Journal*, 14(1):C25–C44, 2011.

[13] C. Schumacher and J. Breitung. Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting*, 24(3):386–398, 2008.

[14] K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, G. Rünstler, K. Ruth, and C. Van Nieuwenhuyze. Short-term forecasting of GDP using large monthly datasets: A pseudo real-time forecast evaluation exercise. *National Bank of Belgium Working Paper*, (133), 2008.

[15] E. Mönch, S. Ng, and S. Potter. Dynamic hierarchical factor models. Technical Report 412, Federal Reserve Bank of New York, 2009.

[16] E. Moench and S. Ng. A hierarchical factor analysis of US housing market dynamics. *The Econometrics Journal*, 14(1):C1–C24, 2011.

[17] J.M. Keynes. On a method of statistical business-cycle research. a comment. *The Economic Journal*, 50(197):154–156, 1940.

[18] D.F. Hendry. Econometrics-alchemy or science? *Economica*, 47(188):387–406, 1980.

[19] Statista, the statistics portal. http://www.statista.com/topics/871/online-shopping/, Accessed [2016-05-19].

[20] Universal Postal Union. The postal statistics. http://www.upu.int/en/resources/postal-statistics, Accessed [2016-05-19].

[21] Organisation for Economic Cooperation and Development. *Languages in a Global World: Learning for Better Cultural Understanding*. OECD Publishing (NJ3), 2012.

[22] S. Wright. *Language policy and language planning: From nationalism to globalisation*. Palgrave Macmillan, 2003.

[23] S.S. Mufwene. Colonisation, globalisation, and the future of languages in the twenty-first century. *International Journal on Multicultural Societies*, 4(2):162–193, 2002.

[24] M. Mowforth and I. Munt. *Tourism and sustainability: Development, globalisation and new tourism in the third world*. Routledge, 2015.

[25] C. Cooper and S. Wahab. *Tourism in the Age of Globalisation*. Routledge, 2005.

[26] E. Smeral. The impact of globalization on small and medium enterprises: new challenges for tourism policies in European countries. *Tourism management*, 19(4):371–380, 1998.

[27] J.H. Mittelman and R. Johnston. The globalization of organized crime, the courtesan state, and the corruption of civil society. *Global Governance*, 5(1):103–126, 1999.

[28] B. Bagley. Globalisation and Latin American and Caribbean organised crime. *Global crime*, 6(1):32–53, 2004.

[29] M. Galeotti. *Global crime today: the changing face of organised crime.* Routledge, 2014.

[30] P.J. Buckley and P.N. Ghauri. Globalisation, economic geography and the strategy of multinational enterprises. *Journal of International Business Studies*, 35(2):81–98, 2004.

[31] A.M. Rugman and A. Verbeke. A perspective on regional and global strategies of multinational enterprises. *Journal of International Business Studies*, 35(1):3–18, 2004.

[32] P. Krugman and A.J. Venables. Globalization and the inequality of nations. Technical report, National Bureau of Economic Research, 1995.

[33] R.C. Feenstra and G.H. Hanson. Globalization, outsourcing, and wage inequality. Technical report, National Bureau of Economic Research, 1996.

[34] R.H. Wade. Is globalization reducing poverty and inequality? *World development*, 32(4):567–589, 2004.

[35] C.E.V. Borio and A.J. Filardo. Globalisation and inflation: New cross-country evidence on the global determinants of domestic inflation. 2007.

[36] J. Ihrig, S.B. Kamin, D. Lindner, and J. Marquez. Some simple tests of the globalization and inflation hypothesis. *International Finance*, 13(3):343–375, 2010.

[37] L.M. Ball. Has globalization changed inflation? Technical report, National Bureau of Economic Research, 2006.

[38] J. Villaverde and A. Maza. Globalisation, growth and convergence. *The World Economy*, 34(6):952–971, 2011.

[39] R. Kaplinsky. *Globalisation, industrialisation and sustainable growth: the pursuit of the nth rent.* Number 365. Institute of Development Studies, University of Sussex, 1998.

[40] M.D. Bordo. Globalisation in historical perspective. *Business Economics*, (37):20–29, 2002.

[41] W. Rennen and P. Martens. The globalisation timeline. *Integrated Assessment*, 4(3):137–144, 2003.

[42] M. Albrow and E. King, editors. *Globalisation, Knowledge and Society.* Sage, London, 1990.

[43] W.W. Powell and K. Snellman. The knowledge economy. *Annual Review of Sociology*, 30:199–220, 2004.

[44] OECD. *Is There a New Economy?* OECD, Paris, 2000.

[45] P.H. Schneider. International trade, economic growth and intellectual property rights: A panel data study of developed and developing countries. *Journal of Development Economics*, 78(2):529–547, 2005.

[46] J.-W. Lee. Capital goods imports and long-run growth. *Journal of Development Economics*, 48(1):91–110, 1995.

[47] A. Harrison. Openness and growth: A time-series, cross-country analysis for developing countries. *Journal of Development Economics*, 48(2):419–447, 1996.

[48] A. Smith. *Wealth of Nations.* 1776.

[49] D. Quah and J.E. Rauch. *Openness and the rate of economic growth.* 1990.

[50] D. Weinhold and J. Rauch. Openness, specialization, and productivity growth in less developed countries. Technical report, National Bureau of Economic Research, 1997.

[51] O.C. Herfindahl. *Concentration in the steel industry.* PhD thesis, Columbia University., 1950.

[52] M. Michaely. Exports and growth: an empirical investigation. *Journal of Development Economics*, 4(1):49–53, 1977.

[53] B. Balassa. Exports and economic growth: further evidence. *Journal of development Economics*, 5(2):181–189, 1978.

[54] B. Balassa. Exports, policy choices, and economic growth in developing countries after the 1973 oil shock. *Journal of development economics*, 18(1):23–35, 1985.

[55] J.F. Helliwell and A. Chung. Macroeconomic convergence: International transmission of growth and technical progress. In *International Economic Transactions: Issues in Measurement and Empirical Research*, pages 388–436. University of Chicago Press, 1991.

[56] J.D. Sachs, A. Warner, A. Åslund, and S. Fischer. Economic reform and the process of global integration. *Brookings papers on economic activity*, 1995(1):1–118, 1995.

[57] F. Rodriguez and D. Rodrik. Trade policy and economic growth: a skeptic's guide to the cross-national evidence. In *NBER Macroeconomics Annual 2000, Volume 15*, pages 261–338. MIT Press, 2001.

[58] E.E. Leamer. Measures of openness. In *Trade policy issues and empirical analysis*, pages 145–204. University of Chicago Press, 1988.

[59] E.F. Heckscher and B.G. Ohlin. *Heckscher-Ohlin trade theory.* MIT Press, 1991.

[60] H.C. Wolf. Trade orientation: measurement and consequences. *Estudios de Economia*, 20:49–86, 1993.

[61] J.-W. Lee. International trade, distortions, and long-run economic growth. *Staff Papers-International Monetary Fund*, 40(2):299–328, 1993.

[62] J.A. Frankel and D.H. Romer. Does trade cause growth? *American Economic Review*, 89(3):379–399, 1999.

[63] S. Naghshpour and B.S. Sergi. World trade indicators and a new approach to measure globalisation and countries' openness. *International Journal of Trade and Global Markets*, 2(1), 2009.

[64] R. Findlay and K.H. O'Rourke. Commodity market integration, 1500–2000. In *Globalization in historical perspective*, pages 13–64. University of Chicago Press, 2003.

[65] R.J. Barro. Economic growth in a cross section of countries. Technical report, National Bureau of Economic Research, 1989.

[66] D. Dollar. Outward-oriented developing economies really do grow more rapidly: evidence from 95 LDCs, 1976–1985. *Economic development and cultural change*, 40(3):523–544, 1992.

[67] R. Summers and A. Heston. The penn world table (mark 5): an expanded set of international comparisons, 1950-1988. *The Quarterly Journal of Economics*, 106(2):327–368, 1991.

[68] J.E. Anderson and J.P. Neary. Measuring the restrictiveness of trade policy. *The World Bank Economic Review*, 8(2):151–169, 1994.

[69] R. Levine and D. Renelt. A sensitivity analysis of cross-country growth regressions. *The American economic review*, 82(4):942–963, 1992.

[70] R.J. Barro and J.-W. Lee. Sources of economic growth. *Carnegie-Rochester Conference Series on Public Policy*, 40:1–46, 1994.

[71] M. Henrekson, J. Torstensson, and R. Torstensson. Growth effects of European integration. *European Economic Review*, 41(8):1537–1557, 1997.

[72] S. Edwards. Openness, productivity and growth: what do we really know? *The economic journal*, 108(447):383–398, 1998.

[73] J. Bryan and T. Sheehy, editors. *Index of economic freedom*. Heritage Foundation, Washington, DC, 1996.

[74] R. Wacziarg. Measuring the dynamic gains from trade. *world bank economic review*, 15(3):393–429, 2001.

[75] J. Randolph. G-index: Globalisation measured. *World Markets Research Centre*, 2001.

[76] P. Martens and D. Zywietz. Rethinking globalization: a modified globalization index. *Journal of International Development*, 18(3), 2006.

[77] A. T. Kearney. Globalisation index. *Foreign Policy*, 2004.

[78] M. Castelli. Measuring... what? notes on some globalization indices. *Globalization*, (5):383–404, 2008.

[79] Ben Lockwood and Michela Redoano. The CSGR globalisation index: An introductory guide. *Centre for the Study of Globalisation and Regionalisation Working Paper*, 155(04), 2005.

[80] A. Dreher. Does globalization affect growth? evidence from a new index of globalization. *Applied Economics*, 38:1091–1110, 2006.

[81] M. Raab, M. Ruland, B. Schönberger, H.-P. Blossfeld, D. Hofäcker, S. Buchholz, and P. Schmelzer. Global index a sociological approach to globalization measurement. *International Sociology*, 23(4):596–631, 2008.

[82] P. Martens and M. Raza. An updated Maastricht globalisation index. *International Centre for Integrated assessment and Sustainable development (ICIS) Working Paper*, 8020, 2008.

[83] L. Figge and P. Martens. Globalisation continues: The Maastricht globalisation index revisited and updated. *Globalizations*, 11(6):875–893, 2014.

[84] P. Vujakovic. How to measure globalization? a new globalization index (ngi). *Atlantic Economic Journal*, 38(2):237–237, 2010.

[85] T.M. Andersen and T.T. Herbertsson. Measuring globalization. Technical Report 817, IZA Discussion paper, 2003.

[86] T.M. Andersen and T.T. Herbertsson. Quantifying globalization. *Applied economics*, (37):1089–1098, 2005.

[87] M. Nishimizu and S. Robinson. Trade policies and productivity change in semi-industrialized countries. *Journal of Development economics*, 16(1–2):177–206, 1984.

[88] A.O. Krueger and B. Tuncer. Growth of factor productivity in Turkish manufacturing industries. *Journal of Development Economics*, 11(3):307–325, 1982.

[89] M. Nishimizu and J. Page. *Trade policy, market orientation and productivity change in industry.* 1991.

[90] D. Rodrik. *Closing technology gaps: does trade liberalization really help.* 1992.

[91] T. Chen and D. Tang. Comparing technical efficiency between import-substitution-oriented and export-oriented foreign firms in a developing economy. *Journal of Development Economics*, 26(2):277–289, 1987.

[92] H. Handoussa, M. Nishimizu, and J.M. Page. Productivity change in Egyptian public sector industries after "the opening", 1973–1979. *Journal of Development Economics*, 20(1):53–73, 1986.

[93] G.S. Olley and A. Pakes. The dynamics of productivity in the telecommunications equipment industry. Technical report, National Bureau of Economic Research, 1992.

[94] A.B. Bernard and J.B. Jensen. Why some firms export. *Review of Economics and Statistics*, 86(2):561–569, 2004.

[95] J.R. Tybout. Linking trade and productivity: New research directions. *The World Bank Economic Review*, 6(2):189–211, 1992.

[96] E. Helpman, M.J. Melitz, and S.R. Yeaple. Export versus FDI. Technical Report 9439, National Bureau of Economic Research, 2003.

[97] A. Young. Learning by doing and the dynamic effects of international trade. Technical report, National Bureau of Economic Research, 1991.

[98] D. Rodrik. Why do more open economies have bigger governments? Technical report, National Bureau of Economic Research, 1996.

[99] E.E. Leamer. What's a recession, anyway? Technical report, National Bureau of Economic Research, 2008.

[100] National Bureau of Economic Research. The NBER's business cycle dating committee. http://www.nber.org/cycles/recessions.html, Accessed [2016-04-29].

[101] J.A. Schumpeter. *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*, volume 55. Transaction publishers, 1934.

[102] A.F. Burns and W.C. Mitchell. *Measuring Business Cycles.* New York: NBER, 1946.

[103] J. Shiskin and G.H. Moore. Composite indexes of leading, coinciding and lagging indicators 1948–67. Supplement to nber report one, NBER, 1968. 1–9.

[104] The Conference Board. *The handbook of cyclical indicators*, 2001. http://www.conference-board.org/pdf_free/economics/bci/BCI-Handbook.pdf, Accessed [2014-11-07].

[105] G. Gyomai and E. Guidetti. OECD system of composite leading indicators, 2012. Accessed [2014-11-09].

[106] V. Zarnowitz and C. Boschan. Cyclical indicators: an evaluation and new leading indexes. *Business Conditions Digest*, 5:5–22, 1975.

[107] The Conference Board. Description of components. https://www.conference-board.org/data/bci.cfm, Accessed [2016-04-07].

[108] M. Chauvet. An econometric characterization of business cycle with factor structure and regime switching. *International Economic Review*, (39):969–996, 1998.

[109] A.J. Auerbach. The index of leading indicators: "measurement without theory, thirty-five years later". *The Review of Economics and Statistics*, (64):589–595, 1982.

[110] J.H. Stock and M.W. Watson. New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, 4:351–409, 1989.

[111] J.H. Stock and M.W. Watson. *Leading Economic Indicators: New Approaches and Forecasting Records.* Number A Probability Model of the Coincident Economic Indicators (Chapter 4). Cambridge, MA: Cambridge University Press, 1991.

[112] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[113] F.X. Diebold and G.D. Rudebusch. Measuring business cycles: A modern perspective. *The Review of Economics and Statistics*, (78):67–77, 1996.

[114] R.H. McGuckin, A. Ozyildirim, and V. Zarnowitz. The composite index of leading economic indicators: how to make it more timely. Working Paper 8430, National Bureau of Economic Research, 2001.

[115] A. Estrella and F.S. Mishkin. Predicting U.S. recessions: Financial variables as leading indicators. *The Review of Economics and Statistics*, 80(1):45–61, 1998.

[116] R.S. Mariano and Y. Murasawa. A new coincident index of business cycles based on monthly and quarterly series. Working paper 14, Penn Institute for Economic Research, 2002.

[117] C. Otrok, C.H. Whiteman, and University of Iowa. Department of Economics. *Bayesian Leading Indicators: Measuring and Predicting Economic Conditions in Iowa.* Working paper series (University of Iowa. Department of Economics). Department of Economics, College of Business Administration, University of Iowa, 1996.

[118] M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic-factor model: Identification and estimation. *Review of Economics and Statistics*, 82(4):540–554, 2000.

[119] I.T. Jolliffe. *Principal Component Analysis, Second Edition.* Springer-Verlag, 1986, 2002.

[120] K. Pearson. On lines and planes of closest fit to systems of points in space. *Pholosophical Magazine*, 2(11):559–572, 1901.

[121] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 6(24):417–441, 1933.

[122] G. Petris, S. Petrone, and P. Campagnoli. *Dynamic linear models*. Springer, 2009.

[123] A.F. Zuur, R.J. Fryer, I.T. Jolliffe, R. Dekker, and J.J. Beukema. Estimating common trends in multivariate time series using dynamic factor analysis. *Environmetrics*, 14(7):665–685, 2003.

[124] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[125] W.K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[126] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-6(6):721–741, 1984.

[127] A.E. Gelfand and A.F.M. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409, 1990.

[128] G.O. Roberts and A.F.M. Smith. Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. *Stochastic processes and their applications*, 49(2):207–216, 1994.

[129] C.A. Sims and T.J. Sargent. Business cycle modeling without pretending to have too much a priori economic theory. *New methods in business cycle research*, 1:145–168, 1977.

[130] J.H. Stock and M.W. Watson. Diffusion indexes. Technical report, National Bureau of Economic Research, 1998.

[131] J.H. Stock and M.W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20(2):147–162, 2002.

[132] J. Bai and S. Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221, 2002.

[133] J. Bai and S. Ng. Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2):304–317, 2008.

[134] D.V. Lindley and A.F.M. Smith. Bayes estimates for the linear model. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(1):1–41, 1972.

[135] D. Gamerman and H.S. Migon. Dynamic hierarchical models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(3):629–642, 1993.

[136] A. Reklaitė. Coincident, leading and recession indexes for the Lithuanian economy. *The Baltic Journal of Economics*, 11(1):91–107, 2011.

[137] G.R. Green and B.A. Beckman. *Survey of Current Business*, volume 73, chapter Business Cycle Indicators: Upcoming Revision of the Composite Indexes, pages 44–51. 1993.

[138] T.M. Crone and A. Clayton-Matthews. Consistent economic indexes for the 50 states. *Review of Economics and Statistics*, 87(4):593–603, 2005.

[139] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *The Annals of statistics*, 32(2):407–499, 2004.

[140] A. Reklaitė. Measuring foreign impact: leading index construction using hierarchical dynamic factor model. *Baltic Journal of Economics*, 16(1):21–32, 2016.

[141] C.K. Carter and R. Kohn. Gibbs sampling for state space models. *Biometrica*, (81):541–553, 1994.

[142] G. Petris. An R package for dynamic linear models. *Journal of Statistical Software*, 36, 2010.

[143] M. J. Faddy and M. C. Jones. Semiparametric smoothing for discrete data. *Biometrika*, 85(1):131–138, 1998.

[144] Marijus Radavičius and Jurgita Židanavičiūtė. Semiparametric smoothing of sparse contingency tables. *Journal of Statistical Planning and Inference*, 139(11):3900–3907, 2009.

[145] S.J. Sheather and M.C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(3):683–690, 1991.