STATISTIKA

STATISTICS AND ECONOMY JOURNAL

VOL. **94** (2) 2014

CZECH STATISTICAL OFFICE

CONTENTS

ANALYSES

METHODOLOGY

INFORMATION

About Statistika

The journal of Statistika has been published by the Czech Statistical Office since 1964. Its aim is to create a platform enabling national statistical and research institutions to present the progress and results of complex analyses in the economic, environmental, and social spheres. Its mission is to promote the official statistics as a tool supporting the decision making at the level of international organizations, central and local authorities, as well as businesses. We contribute to the world debate and efforts in strengthening the bridge between theory and practice of the official statistics. Statistika is a professional double-blind peer reviewed journal included (since 2008) in the List of Czech non-impact peer-reviewed periodicals (updated in 2013). Since 2011 Statistika has been published quarterly in English only.

Publisher

The Czech Statistical Office is an official national statistical institution of the Czech Republic. The Office main goal, as the coordinator of the State Statistical Service, consists in the acquisition of data and the subsequent production of statistical information on social, economic, demographic, and environmental development of the state. Based on the data acquired, the Czech Statistical Office produces a reliable and consistent image of the current society and its developments satisfying various needs of potential users.

Contact us

Journal of Statistika | Czech Statistical Office | Na padesátém 81 | 100 82 Prague 10 | Czech Republic
e-mail: statistika.journal@czso.cz | web: www.czso.cz/statistika_journal

Slowly, we are growing together – European Economic Policy and Statistics[1]

Aurel Schubert[2] | *European Central Bank, Frankfurt am Main, Germany*
Luis M. Serna[3] | *European Central Bank, Frankfurt am Main, Germany*

Abstract

In the last 20 years statistical data has become vastly more important for economic policy in Europe. Whereas economic statistics once played a role in relatively marginal areas of European policy, the establishment of the macroeconomic convergence criteria for joining Economic and Monetary Union in the Maastricht Treaty in 1992–1993 sparked a quantum leap. Questions of comparability and harmonisation suddenly became increasingly relevant. The Stability and Growth Pact then made the calculation of the budget deficit and government debt even more important, including the measurement of GDP as denominator for the respective ratios. With the outbreak of the second Greek crisis in 2009–10 and the flaws that emerged in the quality of Greek economic statistics, statistical questions were suddenly at the centre of international media and political interest. At the same time the financial and economic crisis brought to the fore severe economic imbalances, both between European countries and within European countries. In order to prevent similar imbalances in the future, the EU has developed and adopted the "macroeconomic imbalance procedure", in which currently eleven macroeconomic indicators are used for on-going surveillance of countries ("alert mechanism"). Thus more economic statistics have gained an important political function, particularly since sanctions can even be imposed on the basis of them. In parallel with this, the new European Supervisory Authorities use "dashboards" i.e. a range of statistics that are regularly watched and are intended to function as early warning indicators. The paper takes a look at this move towards more "evidence-based policy making" and its implications for European statistics and statisticians and discusses the related challenges, paying particular attention to the role of the European Central Bank and its specific data needs.

Keywords

Statistics, policy-making, European Central Bank, European System of Central Banks, Eurostat

JEL code

E58, E61, E02

[1] This paper reflects the personal views of the authors and not necessarily those of the ECB or the ESCB. An earlier version of it was presented at the Statistische Woche 2012 in Vienna, Austria, on 21 September 2012.

[2] Directorate General Statistics, European Central Bank, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany. E-mail: aurel.schubert@ecb.int.

[3] Directorate General Statistics, European Central Bank, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany. E-mail: luis.serna@ecb.int.

INTRODUCTION

The role and importance of statistics have varied over time in the European Union's analyses, decisions and communication. Over the years, the removal of barriers to trade between Member States in order to turn the "common market" into a genuine single market in which goods, services, people and capital can move freely has required countless policy decisions not always based on firm statistical information.

The ever growing need for intensified coordination of economic policies of Member States and various Community policies, such as EU regional policies, has gradually led to a parallel increase in the role played by statistics, in a process that can be described as "evidence-based" policy-making. This has been good news for statisticians, who have received increased attention from policy-makers and higher public recognition for their work.

Parallel to that, some of the vast array of economic and financial statistics and indicators available received particular attention when they were used for European administrative purposes. This was the case of the measurement of EU Member States' gross national income at market prices, which has been used in the calculation of the Member States' contributions to the funding of the European Union. Since 1989, European Council legal acts[4] have prescribed that this indicator is to be compiled in application of the European system of national accounts in force (ESA) and provided further details for its compilation.

This example illustrates that in the European context, the simple selection of a statistical indicator for policy or administrative use entails that its comparability and reliability must be reinforced beyond the applicable internationally agreed statistical standards, such as the System of National Accounts (SNA).

For most countries in the world, the SNA is only a methodological manual to which they voluntarily adhere. In Europe, its equivalent, the European System of Accounts, is a legal act.

1 THE 1992 TREATY OF MAASTRICHT: A QUANTUM LEAP FOR STATISTICS IN THE EUROPEAN UNION

The challenges associated with the growing tendency of using statistics for EU economic policies experienced a quantum leap more than 20 years ago with the adoption of the Treaty on European Union signed in Maastricht in February 1992. Three particular elements of this Treaty conferred to statistics a prominent role in the preparatory work for the Economic and Monetary Union (EMU) and the launch of the euro as a single currency:

The first element was the creation of the European Monetary Institute (EMI) and then the European Central Bank, which was entrusted, among other things, with the task of conducting the monetary policy for the single currency, the euro. Furthermore, the Treaty gave powers to the ECB, assisted by the national central banks, to collect the necessary information to support the conduct of its monetary policy, and assigned other tasks to the European System of Central Banks (ESCB), including those related to financial stability.

The second element was the specification of economic convergence criteria for euro adoption, which were based on a set of statistical indicators comprising the Harmonised Index of Consumer Prices (HICP), government deficit and debt, the exchange rate, long-term interest rates and an array of other statistical data to further assess the sustainability of the convergence achieved.

A third element related to the need to maintain and enforce fiscal discipline within EMU after entry, which led to the adoption in 1997 of the Stability and Growth Pact (SGP) further specifying the needs to adhere to quantitative ceilings for government deficits and debt as a percentage of GDP, with the threat of sanctions in case of breaches.

[4] See Council Directive 89/130/EEC, Euratom of 13 February 1989 on the harmonization of the compilation of gross national product at market prices and Council Regulation (EC, Euratom) No 1287/2003 of 15 July 2003 on the harmonisation of gross national income at market prices.

As a result of these three elements embedded in the Maastricht Treaty, quality-related issues such as the availability, comparability and harmonisation of statistics have become increasingly relevant.

2 THE ROLE OF STATISTICS IN THE EUROPEAN ECONOMIC AND MONETARY UNION

The Treaty on European Union specified as one of the tasks of the European Monetary Institute (EMI), the forerunner of the ECB, the preparation of statistics for Stage Three of EMU and, specifically, the promotion of the harmonisation of statistics to the extent necessary. The EMI first released a comprehensive statement of statistical requirements in July 1996.[5]

Moreover, in October 1998, the ECB defined price stability as "a year-on-year increase in the Harmonised Index of Consumer Prices (HICP) for the euro area of below 2%" and added that price stability was "to be maintained over the medium term".[6] Furthermore, the ECB announced its monetary policy strategy based on a comprehensive analysis of the threats to price stability, comprising an "economic analysis" and a "monetary analysis".[7] The "economic analysis" focuses on real economic activity and financial conditions in the economy and is aimed at assessing the short to medium-term determinants of price developments. The economic and financial statistics supporting the analysis include prices and costs statistics, main aggregates of national accounts, government finance statistics, short-term business and labour indicators, exchange rates, the balance of payments of the euro area, financial market statistics and the financial balance sheets of euro area sectors.

The "monetary analysis" focuses on money and liquidity and serves to cross-check, from a medium to long-term perspective, the indications resulting from the "economic analysis". To signal its commitment to monetary analysis, the ECB sets at 4.5% the reference value for the annual growth rate of the broad monetary aggregate M3. This reference value is not a monetary target, but a benchmark for analysis of the information content. The statistics supporting this analysis include, for example, the detailed consolidated balance sheets of euro area banks, in particular the monetary aggregates and counterparts, interest rates on banks' retail loans and deposits, the balance sheets of other monetary and financial institutions, such as investment funds, securities issues statistics and the financial balance sheets of the non-financial sectors.

The need to monitor and cross-check all the relevant statistical information that could influence price stability under the ECB's two-pillar approach resulted in a high demand for harmonised euro area statistics, which were scarce at that time.

2.1 The division of statistical responsibilities between the ECB and Eurostat

The Maastricht Treaty assigned the responsibility for providing official economic and financial statistics to both the ECB, assisted by the national central banks, and the European Commission, which in practice means Eurostat. The division of statistical responsibilities to avoid double work and gaps was established in a memorandum of understanding (MoU) signed in 1995 and amended in 2003.[8]

Under the MoU, the ECB is responsible for monetary and financial statistics. Responsibility for external statistics (balance of payments and international investment position) and for the euro area accounts is shared. Eurostat is responsible for general economic statistics. The latter include the Harmonised Index of Consumer Prices (HICP), which was chosen by the ECB as its benchmark for price stability.

At the starting point of the preparation process there were significant gaps and weaknesses. Where national statistics were available, the underlying concepts, definitions and classifications lacked standardisation.

5 Implementation Package (Statistical Requirements for Stage Three of Monetary Union), EMI, July 1996.

6 The ECB's Governing Council confirmed this definition in May 2003 and clarified that "in the pursuit of price stability, it aims to maintain inflation rates below but close to 2% over the medium term".

7 See Chapter 3.5 of *The monetary policy of the ECB*, ECB, May 2011.

8 See <http://www.ecb.europa.eu/pub/pdf/other/mouecbeurostaten.pdf>.

Describing in detail the extraordinary and cooperative effort made over the years by the European System of Central Banks (ESCB) on one side and Eurostat and the national statistical institutes (NSIs) on the other, is clearly out of the scope of this paper. Moreover, the importance of statistics for monetary policy is described in what follows from the ECB and ESCB viewpoint, which does not mean that the work undertaken by Eurostat and the NSIs in macro-economic statistics has not been equally impressive and important.

2.2 Areas of statistics either under ECB sole responsibility or shared responsibility with Eurostat

The ECB compiled statistics for the euro area using data supplied by the National Central Banks (NCBs) of the euro area, which in turn received the data from reporting agents, and in some limited cases, from the National Statistical Institutes.

By the time the euro was launched in 1999 the initial set of statistics available for the euro area comprised only the bare essentials: harmonised balance sheets of monetary and financial institutions to enable calculation of monetary aggregates and counterparts to money, a limited amount of data on non-harmonised retail interest rates, financial market information acquired from commercial data providers, and key balance of payments statistics. Additionally, annual government finance statistics and some limited annual data on saving, investment and financing were available.

Later in 1999, statistics on securities issues by euro area governments and financial and non-financial corporations were produced. In the subsequent years work continued in this area[9], based on the prioritised user needs. More detailed breakdowns of instruments, maturities and counterpart sectors of Monetary and Financial Institutions' (MFIs) balance sheet statistics were added in 2003. At the same time, harmonised MFI interest rates statistics were introduced, covering a breakdown of both retail deposits and retail loans by maturity and purpose. The ECB took over the release of the daily yield curves for euro area central government bonds from Eurostat. Continuing work in this area, investment fund statistics were released in December 2009, and new harmonised statistics on Financial Vehicle Corporations engaged in securitisation and bank securitisation were published as of June 2011. At the same time, not yet fully harmonised statistics on Insurance Corporations and Pension Funds were published as of June 2011 and work towards their total harmonisation to better serve user needs is on-going. As a result, a large part of the so-called "shadow banking" sector is statistically well covered in the euro area.

Concerning external statistics, the ECB has gradually enhanced the initial set of balance of payments data (b.o.p.) for the euro area by providing more breakdowns, by showing debits and credits separately and by offering a geographical breakdown of major counterparts (e.g. the United States, the United Kingdom, EU countries outside the euro area, Japan and China). The ECB now also produces and publishes a quarterly international investment position (i.i.p.) as well as a breakdown of the changes in the i.i.p. for the euro area.

Moreover, statistics on the nominal and real effective exchange rate and on the international role of the euro have been made available, and have subsequently been supplemented by monthly harmonised national competitiveness indicators based on consumer price indices.

The wide range of statistics available to the ECB to help it fulfil its duties was integrated in June 2007 into quarterly euro area economic and financial accounts by institutional sector. These accounts, compiled together with Eurostat, provide a comprehensive and coherent overview of euro area financial and economic developments. They also show the interrelations between the different sectors in the euro area (households, corporations and general government) and between them and the rest of the world. The full integration and almost complete consistency of these accounts as well as their joint, simultaneous compilation every quarter by two institutions (the ECB and Eurostat) was a major achievement.

9 See BULL, P. *The development of statistics for economic and monetary union.* ECB, June 2004.

Furthermore, many statistics and statistical indicators have also been compiled to assess financial market developments, financial integration in Europe, financial stability overall and within the EU banking sector, and the development of payments, payment infrastructures and securities trading, clearing and settlement.

2.3 Areas of statistics under Eurostat's sole responsibility

By the start of Economic and Monetary Union, Eurostat, in collaboration with the national statistical institutes, had developed the HICP as a harmonised price index as well as other statistics on prices, costs, labour markets and other economic developments. Limited national accounts data for the euro area were available. However, the timeliness of these statistics was not satisfactory for monetary policy purposes. Over time, the timeliness of the relevant euro area statistics provided by Eurostat has improved significantly, following in particular the adoption of the Action Plan for EMU Statistics (2000) by the ECOFIN Council and the establishment of a list of monthly and quarterly Principal European Economic Indicators (PEEIs) in 2002. As an example, timely flash estimates for the HICP and for the quarterly GDP volume changes were important achievements and these data feed into the monetary policy and economic analyses. In addition, the range of available government finance statistics, both annual and quarterly, has significantly expanded. Methodological standards have been further improved in all areas.

However, as described in the 2013 Economic and Financial Committee Status Report on Information Requirements in EMU, further work is still required in terms of new statistics and the frequency and/or quality of statistics available. Improvements are still being made, mainly in the area of services, labour markets (integrating labour market statistics into the national accounts to serve growth and productivity analyses) and housing markets. Further timeliness and other quality improvements are also needed for some other statistics.

2.4 The essential role of cooperation between the ESCB and the ESS

The impressive effort devoted to developing new statistics and to gradually improving them over the years has benefited from close cooperation between the European System of Central Banks (ESCB) on one side and Eurostat and the national statistical institutes on the other (i.e. the European Statistical System (ESS)), notably in the context of the Committee on Monetary, Financial and Balance of Payments Statistics (CMFB). For more than two decades[10] the CMFB has been an important forum for mutual exchange of statistical expertise and has contributed to enhance the collection, compilation and access to high-quality EU and euro area economic and financial statistics. Moreover, the advisory role of the CMFB has been of key importance in the context of the statistics used for the application of the Excessive Deficit Procedure (EDP).

2.5 Fulfilling ECB data needs: a constant challenge for ESCB and ESS statisticians

As a result of the continued effort briefly described in this paper, European statistics produced by both the ESCB and the ESS have improved remarkably over the years and have enabled the ECB to fulfil its prime responsibility of conducting the monetary policy of the euro area and the various tasks entrusted to it in the Treaty.

Ensuring over time that such an array of statistics are of high quality and up-to-date has been a constant challenge that ESCB and ESS statisticians have successfully addressed amid continuous financial

---

[10] To celebrate its 20th anniversary, the Committee on Monetary, Financial and Balance of Payments statistics published the book, *Promoting excellence in European statistics. CMFB 20 years,* <http://www.cmfb.org/pdf/2011-11-25%20CMFB_Promoting%20Excellence%20in%20European%20Statistics.pdf>.

innovation, euro area and EU enlargements, the challenges posed by globalisation and the increasing demand for statistics for the purposes of resolving the current financial and economic crisis.

3 DATA DEMANDS BEYOND AGGREGATES AND AVERAGES

European policy-makers' demands to close obvious data gaps evidenced by the crisis have increased exponentially. At the same time, they have come to demand more detailed data beyond the traditional aggregate approach to statistics, requesting information on distributions around the averages, and, in addition to the euro area aggregates, also country level information. Especially, the early identification of economic vulnerabilities within Europe or the euro area required such detailed information.

Thus, in areas in which no comparable and timely data existed, such as the financial conditions faced by small and medium enterprises (SMEs), the ECB had to go beyond the traditional census type statistics and develop, in collaboration with the European Commission, a survey on the access to finance of SMEs in the European Union. Since mid-2009 this survey has been conducted every two years, while the ECB runs part of the survey every six months for companies in the euro area in order to assess the latest developments in their financing conditions.

In addition, the ECB has implemented, together with 15 Eurosystem NCBs and in close cooperation with a number of national statistical institutes, a survey of household finance and consumption in the euro area whose first results have been published in April 2013. The survey provides micro-level data on households' real and financial assets, liabilities, consumption and saving, income and employment, future pension entitlements, intergenerational transfers and gifts, and attitudes to risk.

Both surveys are vivid examples of how statisticians have responded to the changing information needs of European policy makers.

4 THE IMPORTANT ROLE OF STATISTICS IN CONFRONTING THE FINANCIAL AND ECONOMIC CRISIS

One of the key lessons from the crisis has been that policy-makers need more and more timely quantitative information in order to take good decisions or make recommendations. For the ECB, this has translated into an additional high demand for statistics in the areas of financial stability and macro-prudential supervision, areas in which it has to lend support to the recently created European Systemic Risk Board (ESRB).[11]

The financial crisis which started in 2007 revealed some weaknesses in the governance of the European Economic and Monetary Union, particularly in the field of EU macro- and micro-financial supervision. Moreover, weaknesses were also detected in the governance of the euro area and the EU in relation to the Stability and Growth Pact (SGP) and surveillance of euro area countries' macroeconomic imbalances.

All these weaknesses have in the meantime been addressed by legislative measures which represent the most comprehensive reinforcement of the EU architecture since the launch of the EMU project 20 years ago. Their implementation entails a major step forward for statistics as selected sets of warning indicators (dashboards, scoreboards) play an increasingly important policy role and may trigger recommendations, warnings and in some cases fines. Naturally, besides the implied additional demand for statistics, we can expect an increased public interest in such indicators, also at international level.[12]

[11] The ECB shall ensure a Secretariat, and thereby provide analytical, statistical, logistical and administrative support to the ESRB (Article 2 of Council Regulation (EU) No 1096/2010).

[12] The 12 December 2012 US Federal Reserve policy statement announcing that it will be appropriate to maintain its exceptionally low range for the federal funds rate "at least as long as the unemployment rate remains above 6.5%" is an example of how an indicator may receive increased international attention when it is selected as policy target.

4.1 ESRB Dashboard of indicators for macro-prudential oversight of the financial system

The financial crisis revealed important shortcomings in EU financial supervision, which had failed to anticipate adverse macro-prudential developments and to prevent the accumulation of excessive and systemic risks within the financial system. A proper functioning of EU and global financial systems and the mitigation of threats to them required enhanced consistency between macro and micro-prudential supervision.

To remedy the situation and prevent a future crisis, the EU decided to bring together the actors of financial supervision at national level and at the level of the EU to act as a network. The European System of Financial Supervision (ESFS), which started its work on 1 January 2011, comprises the European Systemic Risk Board (ESRB), which is in charge of macro-prudential oversight, and three micro-prudential supervisory authorities: the European Banking Authority, the European Insurance and Occupational Pensions Authority and the European Securities and Markets Authority.

In its responsibility for macro-prudential oversight across the entire EU financial system, the ESRB is charged with identifying risks to financial stability and, where necessary, with issuing risk warnings and recommendations for action to address such risks.

The ECB plays a prominent role as, in accordance with the legislation, it provides "analytical, statistical, administrative and logistical support to the ESRB, also drawing on technical advice from national central banks and supervisors".

In its duty to support the ESRB in the area of statistics, the ECB helps to ensure that appropriate and reliable information is made available for the ESRB to perform its duties, while preserving the confidentiality of that information as is legally required.

In accordance with the ESRB Regulation, the ECB has developed, together with the ESRB Secretariat, a Risk Dashboard comprising about 45 relevant indicators to identify and assess potential systemic risks. Those systemic risks include risks of disruption to financial services caused by a significant impairment of all or parts of the EU financial system that could potentially have serious negative consequences for the internal market and the real economy. Any type of financial institution and intermediary, market, infrastructure and instrument has the potential to be systemically significant.

The ESRB Risk Dashboard, which was published on 20 September 2012 for the first time, is divided into six categories: inter-linkages and composite measures of systemic risk, macro risk, credit risk, funding and liquidity, market risk and profitability and solvency. The data sources are the ECB, Eurostat and commercial data providers.

In addition to developing the indicators, both in terms of choosing the right type of indicator and structuring the Dashboard effectively, the work of statisticians includes assessing the quality of each indicator chosen, partly through back-testing. Annual updates of the list of indicators will ensure that the Dashboard remains a relevant tool for identifying and measuring systemic risk.[13]

4.2 Fiscal surveillance in the amended Stability and Growth Pact

The economic and financial crisis has exacerbated the pressure on the public finances of EU Member States. Government finance statistics, which are a key element in supporting fiscal surveillance under the Stability and Growth Pact, received renewed attention when the fiscal imbalances of several European countries led to increases in their sovereign risks.

Under the reinforced SGP adopted in December 2011, financial sanctions apply to euro area Member States that do not take adequate action to bring their budget deficits below 3% of GDP within the agreed timeframe. The financial sanctions are imposed unless a "reverse qualified majority" of Mem-

[13] The three European Supervisory Authorities have also developed their respective dashboards, although they have been not (yet) published.

ber States vote against it, which makes the enforcement of the rules more automatic, dissuasive and credible.

Moreover, under the amended SGP the debt criterion of the Treaty, which established a general government debt benchmark, namely a 60% debt-to-GDP ratio, has received more emphasis and must be respected after a transitional period. After taking into account all the relevant factors and the impact of the economic cycle, if the gap between its debt level and the 60% reference is not reduced by $1/20^{th}$ annually (on average over three years), the Member State concerned will be subject to the excessive deficit procedure, even if its deficit is below 3%.

The above-mentioned corrective measures are further complemented by preventive measures based on country-specific medium-term objectives, compliance with which will be closely monitored based on statistics, with the possibility of imposing financial sanctions in the form of non-interest-bearing deposits and fines.

In this context, ensuring that the data used for EU fiscal surveillance meets the necessary statistical quality standards is of utmost importance. The Commission (Eurostat) was given this responsibility within the framework of the Excessive Deficit Procedure (EDP) at the time of its implementation in 1994. However, Eurostat does not compile directly government data for the Member States but depends greatly on the data compiled and reported by them, as well as on the administrative ability, goodwill and cooperation of the respective national statistical authorities. On specific occasions, this framework has resulted in data misreporting, which has led to a number of measures being taken over the years to strengthen the EU governance of fiscal statistics. Among them, Council Regulation (EU) No 679/2010 has granted Eurostat new competences for regularly monitoring and verifying public finance data, which it will exercise by conducting more in-depth dialogue visits to Member States and by extending such visits to public entities supplying upstream public finance data to the NSIs.[14]

The traditional role of the ECB in monitoring government finance statistics, e.g. via its participation in the CMFB, has even increased with the ECB role in country missions (Troika), the ECB's bond purchases under the Securities Markets Program and possible Outright Monetary Transactions. Access to detailed, timely and higher government finance data is absolutely crucial for the ECB.[15]

4.3 Scoreboard of indicators for macroeconomic imbalances surveillance

Similarly to fiscal surveillance, the new mechanism for identifying and correcting competitiveness gaps and major macroeconomic imbalances relies heavily on statistical information.

The Macroeconomic Imbalances Procedure (MIP) is a new surveillance and enforcement mechanism based on Article 121.6 of the Treaty on the functioning of the European Union and relies on the following main elements:

- preventive and corrective action: before the imbalances become large this procedure allows the Commission and the Council to adopt preventive recommendations at an early stage. In more serious cases, there is also a corrective limb, under which an excessive imbalance procedure can be opened for a Member State;
- rigorous enforcement consisting of a two-step approach whereby an interest-bearing deposit can be imposed after one failure to comply with the recommended corrective action. After a second compliance failure, this interest-bearing deposit can be converted into a fine (up to 0.1% of GDP).

[14] Section 5 elaborates further on the importance of following sound quality principles for developing, producing and disseminating statistics in an independent manner and free from any political pressure.

[15] This essential role of government finance statistics for the ECB is the rationale for having a dedicated ESCB working group for government finance statistics (WG GFS).

The semi-automatic decision-making process uses reverse qualified majority voting to decide on sanctions, making it very difficult for Member States to form a blocking majority; and

- an early warning system: an alert system is established based on an economic reading of a scoreboard consisting of a set of statistical indicators covering the major sources of macroeconomic imbalances. The scoreboard contains thresholds for the indicators which trigger further in-depth analyses to determine the gravity of potential imbalances, with the help of a broader set of indicators.

The early involvement of statisticians from both the ESS and the E(S)CB in the discussions concerning the statistical aspects has ensured that the ten indicators of the scoreboard launched on 14 February 2012 are relevant, simple, measurable, available in good time and based on a solid statistical methodological framework. Moreover, this work is being completed by the development of quality profiles for the indicators.

The composition of the scoreboard indicators may evolve over time to ensure that it remains relevant. To this end, the revised MIP scoreboard released on 28 November 2012 incorporated an additional indicator related to the financial sector, namely the total liabilities of the financial sector (see Annex 1). Acknowledging the importance of timely statistics of the highest quality for the credibility of the MIP procedure, the ECOFIN, in its conclusions of 8 November 2011 and 13 November 2012, invited the ESS and the ESCB "to work together on improving the underlying statistics and to ensure comparability".[16]

5 QUALITY ISSUES AND THE GREEK CASE

With the increased use of statistics by European policy-makers, the quality of European official statistics has become a very important issue.

In a democratic society, official statistics are one of the cornerstones of good government and public confidence in good government. It is therefore fundamental to ensure the highest possible quality standards in the compilation and dissemination of statistics.

In the early 1990s, the public's trust in official statistics in various countries, particularly in Central Europe and the former Soviet Union, was impaired. To remedy the situation, the Conference of European Statisticians adopted the Fundamental Principles of Official Statistics (1992), which were subsequently endorsed by the United Nations Statistical Commission (1994).[17]

As Principle one states: *"Official statistics provide an indispensable element in the information system of a democratic society, serving the government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information."*

5.1 Quality Framework in the European Statistical System (ESS)

Enhanced surveillance of fiscal, macroeconomic and macro-prudential policies must rely on statistical information produced under robust quality management. At the same time, the quality of the data is key for any credible policy process.

During the economic and financial crisis the insufficient quality of fiscal data provided to Eurostat by Member States jeopardised the credibility of the entire fiscal surveillance framework of the EDP once more.

---

[16] Refer to the draft Regulation of the European Parliament and of the Council on the provision and quality of statistics for the macroeconomic imbalances procedure <http://www.ipex.eu/IPEXL-WEB/dossier/files/download/082dbcc53eea9c03013f1dcf7c0816e7.do> and the ECB opinion on it <http://www.ecb.europa.eu/ecb/legal/pdf/en_con_2013_72_f_sign.pdf>.

[17] See <http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx>.

The triggering element was a new "Greek case", i.e. the renewed problems in the Greek fiscal statistics,[18] which pointed out that the measures taken after the 2004 Greek misreporting of the Excessive Deficit Procedure data required further reinforcement. The governance of the European Statistical System (ESS) had been improved, in particular with the adoption of the Code of Practice in 2005, but its implementation and monitoring relied to a large extent on a self-regulatory approach (self-assessments, peer reviews and national implementation plans). In 2009, the situation changed when a newly created body, the European Statistical Governance Advisory Board (ESGAB),[19] started work in order to provide an independent overview of the implementation of the Code by Eurostat and the NSIs and the Regulation on European Statistics entered into force.

In its first report of November 2009, the ESGAB included some general recommendations regarding the institutional set-up in reaction to the first information about the new Greek case: "An appropriate institutional framework is crucial in order to safeguard the professional independence of statistical authorities. Suspicions of interventions affecting the data produced need to be further investigated. Moreover, the procedures for the appointment and dismissal of Heads of National Statistical Institutes (NSIs) should be transparent and kept separate from political mandates." The report also stresses that "a stronger commitment from top management in the statistical offices and a stronger adherence to common quality standards at the level of the ESS will be of essence".

Subsequently, the financial crisis evolved into a fiscal crisis in Europe which required further measures to address the remaining weaknesses in the governance framework of the ESS, including granting Eurostat new competences on EDP matters. These weaknesses were described in the Communication from the Commission to the European Parliament and the Council of 15 April 2011 entitled "Towards robust quality management for European Statistics".[20]

In this context, the Code of Practice was revised in September 2011 in order to distinguish between the principles to be implemented by ESS members and the principles relating to the institutional environment that are to be implemented by Member State governments.

Moreover, the Regulation on European Statistics[21] is currently under revision with a view to clarifying, among other things, that the principle of professional independence of NSIs applies unconditionally. Statistics must indeed be developed, produced and disseminated in an independent manner, free of any pressure from political or interest groups or from EU or national authorities, and existing institutional frameworks must not be allowed to restrict this principle.

5.2 Quality Framework in the European System of Central Banks (ESCB)

Credibility has always been crucial for central banks, and therefore, for the statistics they produce.

Since the establishment of the ECB, adherence to high-quality standards has been considered a key determinant in maintaining the public's confidence in the ECB statistics upon which policy decisions are based. Currently, this is also of upmost importance in view of the ECB's provision of statistical support to the European Systemic Risk Board (ESRB).

The ECB statistics follow widely agreed global and European statistical standards such as the System of National Accounts, the European System of Accounts and the Balance of Payments Manual. Moreover, the ECB actively cooperates with the relevant international organisations (Eurostat, IMF, BIS, OECD, UN) to achieve worldwide harmonisation of standards and definitions for economic and financial statistics.

---

18 See the *Report on Greek government deficit and debt statistics.* European Commission, January 2010.

19 See ESGAB's website <http://epp.eurostat.ec.europa.eu/portal/page/portal/esgab/introduction>.

20 See <http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/COM-2011-211_Communication_Quality_Management_EN.pdf>.

21 See <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:087:0164:0173:en:PDF>.

In a clear commitment to high-quality statistical work, the ECB and the NCBs also follow internationally agreed quality standards, such as those formulated in the IMF's Special Data Dissemination Standard and Data Quality Assessment Framework, which are in turn rooted in the UN Fundamental Principles of Official Statistics.

Even though the Treaty provisions concerning the ESCB's independence and accountability also apply to its statistical function, the ESCB collaborates with the European Statistical System (ESS) and respects and applies rigorously the principles laid down in the European Statistics Code of Practice for the National and Community Statistical Authorities.[22]

Thus, the statistical principles underlying the European statistics produced by the ESCB are currently set out in Article 3a of Council Regulation 2533/98 as amended in 2009: "The development, production and dissemination of European statistics by the ESCB shall be governed by the principles of impartiality, objectivity, professional independence, cost effectiveness, statistical confidentiality, minimisation of the reporting burden and high output quality, including reliability and the definitions of these principles shall be adopted, elaborated on and published by the ECB. These principles are similar to the statistical principles of Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics."

All these elements are covered by the recently amended "Public commitment on European statistics by the ESCB", which stresses adherence to high-quality standards when collecting, compiling and disseminating statistics under the ESCB's responsibility.[23]

6 COMMUNICATION ASPECTS

Policy-makers should communicate effectively in order to make their policy decisions accountable, transparent and well understood by the public. There is no doubt that disseminating reliable statistics should be considered an integral part of the communication strategy of any policy-maker and so part of "evidence-based policy making".

As Alexandre Lamfalussy, the first president of the European Monetary Institute, said in 1996, "nothing is more important for monetary policy than good statistics. Statistical information is necessary to decide what policy actions to take, to explain them publicly, and to assess their effect after the event".

Given the independence of the ECB, it is of upmost importance that the public has the possibility to hold the ECB accountable for its policy decisions. A high degree of transparency helps to make the monetary policy more credible and effective.

The ECB publishes in quasi-real time the information on which its Governing Council has based its decisions and its President explains the diagnosis of the situation and actions taken at a press conference. A few weeks later, the ECB's Monthly Bulletin explains the analysis in more detail and provides statistical evidence in an annex. Furthermore, the large array of statistics is published expeditiously on the ECB's website in line with a release calendar.

In this regard, disseminating the statistics associated with ECB policy decisions has proved to be an effective communication tool, a quasi-policy tool.

However, the pre-requisite for statistics to be an effective channel in communication is that policy-makers and statisticians are able to understand each other. Communication between statisticians and policy-makers should be improved because statistical concepts such as those embedded in statistical methodological manuals are sometimes regarded as very technical and are not easily understood by politicians and policy-makers. Therefore, statisticians must make an effort to understand the European

[22] See <http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-32-11-955/EN/KS-32-11-955-EN.PDF>.
[23] See <http://www.ecb.europa.eu/stats/html/pcstats.en.html>.

policy needs, which are usually expressed in "political language" and translate them subsequently into "technical language" to ensure precision on what needs to and can be measured.

In this process, the close cooperation between policy-makers and statisticians at national and European level in the design and selection of the most appropriate indicators for policy purposes has proven to be a key success factor and has enabled to avoid subsequent "clean up" work.

On the other hand, policy-makers should guarantee the scientific and professional independence of statisticians. Experience has demonstrated that the more relevant the statistics are for policy making and policy evaluation, the higher the temptation is for politicians to influence the impartiality of the statistics when they have failed to achieve the policy goals underlying them.

In turn, with the increased public interest in statistical work, statisticians should refrain from entering into "politics" by trying to attract the attention of mass media by interpreting politically the developments they have measured and identified.

Given the European context, preserving the boundaries of the professional independence of statisticians is a key element for producing high-quality statistics and for building trust among Member States, European institutions and the EU citizens which they serve.

CONCLUSIONS

Statistics are increasingly present in all European policy decisions because they provide the evidence for analysis, decision-making and transparent communication.

The decisions taken by European policy-makers to overcome the effects of the economic and financial crisis rely heavily on (new) sets of indicators which try to summarise complex developments. This has increased the policy relevance of statistics, which, besides providing the basis for good decision making, now can trigger automatic action when certain agreed thresholds are exceeded. Furthermore, a lack of corrective action in response to warnings and recommendations may eventually lead to the imposition of fines.

It is therefore fundamental to ensure that European statistics are of the highest possible standard in order to make them sound and undisputable. If statisticians produce statistics impartially and free from political or any other external pressure, policy-makers will have at their disposal a powerful tool to assess the situation correctly, implement appropriate and credible measures and explain their decisions to the public. Moreover, citizens are entitled to receive objective information on a given situation and to be able to monitor the results of policies.

In difficult times, in which unpopular decisions need to be taken, good quality statistics contribute to building up mutual trust. During the last two decades, enormous efforts have been made in this respect by both systems of European statistics, the ESS as well the ESCB, but ensuring the quality of statistics remains a never ending challenge. As the economies and the policy challenges develop further, the statistics have to evolve too to stay relevant for the policy makers, in order to allow for the right policy answers. The process of "growing together" has to go on.

In this endeavour, adequate resources and close cooperation among statisticians, both academic and official, as well as with policy-makers at national and European level, are key factors for success.

Statistical institutes and central banks have to work closely hand-in-hand to ensure the necessary level of quality and with that, the continuous credibility of European statistics.

References

BULL, P. *Statistics for economic and monetary union – enhancements and new directions, 2003–2012* [online]. ECB, December 2013. <http://www.ecb.europa.eu/pub/pdf/other/statisticsforeuropeanmonetaryunion2013en.pdf>.

BULL, P. *The development of statistics for economic and monetary union* [online]. ECB, June 2004. <https://www.ecb.int/pub/pdf/other/developmentstatisticsemu200406en.pdf>.

CMFB. *Promoting excellence in European statistics. CMFB 20 years* [online]. 2011. <http://www.cmfb.org/pdf/2011-11-25%20CMFB_Promoting%20Excellence%20in%20European%20Statistics.pdf>.

ECB. *Convergence Report 2014* [online]. June 2014. <http://www.ecb.europa.eu/pub/pdf/conrep/cr201406en.pdf>.

ECB. *ECB statistics: an overview* [online]. April 2010. <http://www.ecb.int/pub/pdf/other/ecbstatisticsanoverview2010en.pdf>.

ECB. *Public commitment on European Statistics by the ESCB* [online]. 2012. <http://www.ecb.europa.eu/stats/html/pcstats.en.html>.

ECB. *The ECB Statistics Quality Framework and quality assurance procedures* [online]. <http://www.ecb.int/stats/html/sqf.en.html>.

ECB. *The monetary policy of the ECB* [online]. May 2011. <http://www.ecb.int/pub/pdf/other/monetarypolicy2011en.pdf>.

ECB and EUROSTAT. *Memorandum of Understanding on economic and financial statistics between the Directorate General Statistics of the ECB (DG Statistics) and the Statistical Office of the European Communities (Eurostat)* [online]. March 2003. <http://www.ecb.int/pub/pdf/other/mouecbeurostaten.pdf>.

ECOFIN. *Action Plan on EMU statistical requirements* [online]. September 2000. <http://www.consilium.europa.eu/uedocs/cmsUpload/Commission%20Action%20Plan%20on%20EMU%20Statistical%20Requirements.pdf>.

ECOFIN. *Council conclusions on an early warning scoreboard for the surveillance of macroeconomic imbalances* [online]. 8 November 2011. <http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ecofin/125974.pdf>.

ECOFIN. *Council conclusions on EU statistics, 3198th Council meeting press release (13 November 2012)* [online], pp. 24–28. <http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ecofin/133470.pdf>.

ECOFIN. *Status Report on Information Requirements in EMU* [online]. 15 November 2013. <http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/EFC_REPORT_201313/EN/EFC_REPORT_201313-EN.PDF>.

EMI. *Implementation Package (Statistical Requirements for Stage Three of Monetary Union).* July 1996.

ESGAB. *Fourth annual report by the European Statistical Governance Advisory Board to the European Parliament and the Council* [online]. 2012. <http://epp.eurostat.ec.europa.eu/portal/page/portal/esgab/documents/EN_ESGAB%202012%20report.pdf>.

ESRB. *European Systemic Risk Board's Risk Dashboard (Issue 6)* [online]. January 2014. <http://www.esrb.europa.eu/pub/pdf/dashboard/140128_ESRB_risk_dashboard.pdf?2adc4cae77aecc8e487c462676a6bfb5>.

EUROPEAN COMMISSION. *Report on Greek government deficit and debt statistics* [online]. January 2010. <http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/COM_2010_REPORT_GREEK/EN/COM_2010_REPORT_GREEK-EN.PDF>.

EUROPEAN COMMISSION. *Report on the Alert Mechanism Report 2013* [online]. November 2012. <http://ec.europa.eu/europe2020/pdf/amreport2013_en.pdf>.

EUROPEAN COMMISSION. *Towards robust quality management for European Statistics* [online]. April 2011. <http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/COM-2011-211_Communication_Quality_Management_EN.pdf>.

EUROSTAT and ESS. *The European Statistics Code of Practice for the national and community statistical authorities* [online]. September 2011. <http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-32-11-955/EN/KS-32-11-955-EN.PDF>.

UN. *Fundamental Principles of Official Statistics* [online]. United Nations Statistics Division, 1994. <http://unstats.un.org/unsd/methods/statorg/FP-English.htm>.

ANNEX

Table 1 MIP scoreboard with values for 2012

Year 2012	External Imbalances and Competitiveness							External Imbalances and Competitiveness		Internal Imbalances						
	Current Account Balance as % of GDP		Net International Investment Position as % of GDP	Real Effective Exchange Rate (42 IC – HICP deflators)		Export Market Shares		Nominal ULC		% y-o-y change in Deflated House Prices	Private Sector Credit as % of GDP, consolidated	Private Sector Debt as % of GDP, consolidated	General Government Sector Debt as % of GDP	Unemployment Rate		% y-o-y change in Total Financial Sector Liabilities
	3 years average	*p.m.: level year 2012*		% change (3 years)	*p.m.: % y-o-y change*	% change (5 years)	*p.m.: % y-o-y change*	% change (3 years)	*p.m.: % y-o-y change*					3 years average	*p.m.: level year 2012*	
Thresholds	−4/+6%	-	−35%	±5%&±11%	-	−6%	-	+9%&+12%	-	6%	14%	133%	60%	10%	-	16.5%
BE	−0.4	*−2.0*	48	−4.3	*−2.3*	−14.9	*−5.2*	6.6	*4.1*	−0.2	−1.5	146	100	7.7	*7.6*	−3.9
BG	−0.9	*−1.3*	−80	−4.0	*−2.0*	4.8	*−5.5*	7.4	*−0.5*	−5.3 (p)	2.5	132	19	11.3	*12.3*	10.1
CZ	−3.0	*−2.4*	−50	0.4	*−2.8*	−4.2	*−4.6*	3.9	*3.8*	−3.9	0.6	72	46	7.0	*7.0*	5.4
DK	5.9	*6.0*	38	−7.7	*−2.8*	−18.6	*−4.8*	1.0	*1.6*	−5.1	6.1	239	45	7.5	*7.5*	5.0
DE	6.5	*7.0*	42	−8.9	*−3.2*	−13.1	*−4.6*	3.0	*3.1*	1.8	1.5	107	81	6.2	*5.5*	4.4
EE	0.9	*−1.8*	−54	−3.4	*−0.6*	6.5	*−4.1*	−2.8	*4.2*	3.5	4.7	129	10	13.2	*10.2*	12.9
IE	2.3	*4.4*	−112	−12.2	*−4.3*	−16.3	*−3.3*	−10.4	*0.0*	−11.7	−1.6	306	117	14.4	*14.7*	−0.7
EL	−7.5	*−2.4*	−109	−4.5	*−3.9*	−26.7	*−7.3*	−8.1	*−6.2*	−12.4 (1)	−6.8	129	157	18.2	*24.3*	−3.4
ES	−3.1	*−1.1*	−93	−5.2	*−2.3*	−14.6	*−4.9*	−5.6	*−3.0*	−16.9	−10.5	194	86	22.3	*25.0*	3.3
FR	−1.8	*−2.2*	−21	−7.8	*−3.2*	−14.0	*−6.8*	4.1	*2.1*	−2.3	3.5	141	90	9.9	*10.2*	−0.1
HR	−0.5	*0.0*	−89	−8.3	*−2.6*	−24.7	*−7.4*	0.8	*1.2*	−2.4	−2.1	132	56	13.8	*15.9*	0.9
IT	−2.3	*−0.4*	−25	−6.2	*−1.8*	−23.8	*−5.0*	3.1	*2.3*	−5.4 (p)	−1.0	126	127	9.2	*10.7*	7.1
CY	−6.7	*−6.9*	−82	−5.8	*−1.9*	−26.6	*−9.4*	0.8	*−2.7*	−2.2	10.0	299	87	8.7	*11.9*	−1.9
LV	−0.6	*−2.5*	−67	−8.5	*−1.4*	12.3	*5.4*	−5.8	*3.4*	−0.6	−1 (p)	91.7 (p)	41	16.9	*15.0*	4.1 (p)
LT	−1.3	*−0.2*	−53	−6.7	*−2.0*	29.3	*5.7*	−4.6	*1.9*	−3.2	−0.3	63	41	15.6	*13.4*	−0.3
LU	7.0	*6.6*	169	−2.3	*−1.4*	−18.3	*−4.0*	9.8	*4.7*	2.5	−5.0	317	22	4.8	*5.1*	11.3
HU	0.6	*1.0*	−103	−1.2	*−2.3*	−17.8	*−7.4*	4.4	*2.7*	−9.2	−6.1	131	80	11.0	*10.9*	−8.3
MT	−1.6	*1.6*	25	−7.7	*−2.1*	4.5	*−1.9*	4.9	*3.7*	0.3	−1.6	155	71	6.6	*6.4*	4.1
NL	8.8	*9.4*	47	−6.0	*−1.8*	−12.0	*−3.3*	3.3	*2.8*	−8.7	0.2	219	71	4.7	*5.3*	4.9
AT	2.2	*1.6*	0	−4.7	*−1.7*	−21.2	*−6.3*	4.1	*3.0*	na (2)	2.7	147	74 (3)	4.3	*4.3*	−0.9
PL	−4.6	*−3.7*	−67	1.3	*−2.3*	1.3	*−2.7*	4.4	*2.0*	−5.9 (e)	3.4	75	56	9.8	*10.1*	9.6
PT	−6.5	*−2.0*	−115	−4.0	*−1.5*	−16.0	*−5.3*	−5.3	*−3.1*	−8.6 (p)	−5.4	224	124	13.6	*15.9*	−3.6
RO	−4.4	*−4.4*	−68	−1.9	*−6.0*	5.9	*−7.1*	4.8	*6.5*	−9.2	0.9	73	38	7.2	*7.0*	5.3
SI	1.2	*3.3*	−45	−4.5	*−1.2*	−19.9	*−6.9*	0.4	*0.8*	−8.4	−2.9	114	54	8.1	*8.9*	−0.8
SK	−1.7	*2.2*	−64	−3.2	*0.0*	4.2	*1.5*	0.9	*1.0*	−5.9	3.2	73	52	14.0	*14.0*	2.6
FI	−0.5	*−1.7*	18	−8.3	*−2.7*	−30.8	*−7.1*	4.8	*4.4*	−0.5 (p)	9.0	158	54	8.0	*7.7*	−0.2
SE	6.2	*6.0*	−10	10.1	*−0.8*	−18.8	*−6.0*	0.7	*2.9*	−0.2	1.8	212	38	8.1	*8.0*	4.4
UK	−2.8	*−3.8*	−9	5.8	*4.3*	−19.0	*−1.7*	6.1	*3.0*	−0.9	2.6	179	89	7.9	*7.9*	−4.3

Note: (1) Eurostat estimate based on HPI data from Bank of Greece produced in agreement with ELSTAT. (2) HPI data up until 2011 by Statistics Austria. For 2012, Eurostat estimates a deflated rate of 9.6% based on nonharmonised HPI data by ECB and Central Bank of Austria. (3) Eurostat expressed a reservation on Austrian general government sector debt, see Eurostat press release 152/2013. (4) p = provisional data, e = estimated.

Source: EUROSTAT, DG ECFIN (for the indicators on the REER)

Evaluation of the Ministry of Finance's Forecast History

Petra Vacková[1] | *Ministry of Finance of the Czech Republic, Prague, Czech Republic*

Abstract

This paper evaluates the accuracy of macroeconomic economic forecasts of the Ministry of Finance of the Czech Republic using the average forecasting error, the mean absolute error and Theil's inequality coefficient. The paper analyses the forecast accuracy of the main macroeconomic indicators – real GDP growth, nominal GDP growth, GDP deflator growth, real private consumption growth, average inflation rate, average unemployment rate and current account balance to GDP Ratio. The forecast accuracy is also assessed using the modified Diebold and Mariano test, which compares the accuracy of two forecasts under the null hypothesis that assumes no differences in accuracy. Last but not least, the paper compares the accuracy of the forecasts of the Ministry of Finance to those of the European Commission, Organization for Economic Co-operation and Development and International Monetary Fund.

Keywords

The accuracy of macroeconomic forecasts, the average forecasting error, the mean absolute error, Theil's inequality coefficient, the naïve forecast, modified Diebold-Mariano test

JEL code

C82

INTRODUCTION

This analysis evaluates the forecast accuracy of the macroeconomic forecasts of the Ministry of Finance of the Czech Republic. The first experimental publication summarizing the past and expected future development of basic economic indicators was published by the Ministry of Finance in November 1995. Today, an 18-year history of regular quarterly forecasts provides a high-quality source with which to evaluate their success rate. This can help forecast users to get an idea of how precisely the Ministry of Finance is able to predict the future development of basic macroeconomic indicators across various time horizons.

It is necessary to note that all macroeconomic forecasts are inherently conditioned by adopted assumptions regarding the development of exogenous factors, of which some, for example natural disasters, the development of financial markets, including commodity prices or changes in the political environment outside and inside the Czech Republic, are inherently unpredictable. Other assumptions, for example the impact of structural policy measures, can only be quantified with great difficulty. Another important source of uncertainty is revisions of databases for past periods, concerning in particular those most important indicators of the national accounting system (GDP and its components).

Last but not least, it is necessary to point out the fact that at a time of economic turbulence and financial crises the forecasting of future economic development is considerably more difficult than in a period of stable economic growth.

[1] Ministry of Finance of the Czech Republic, Letenská 15, 118 10 Prague 1, Czech Republic. E-mail: Petra.Vackova@mfcr.cz.

Identifying the impacts of those factors emanating externally and which are completely beyond the control of the forecasting team is, however, difficult (if not impossible) and therefore in accordance with literature I have abstracted away from these facts.

1 DATA

The estimates of the future development of main economic indicators are published in the macroeconomic forecasts of the Ministry of Finance, which has been released quarterly since November 1995. This survey analyses the forecast accuracy for several macroeconomic indicators (real GDP growth, nominal GDP growth, GDP deflator growth, real private consumption growth, average inflation rate, average unemployment rate and current account balance to GDP Ratio).

I have divided the period 1995–2012 into three six-year periods of identical length (1995–2000, 2001–2006 and 2007–2012)[2] in order to be able to evaluate the success rate of forecasts over time. It is necessary to point out that during the evaluated period some major changes have occurred in the Czech economy, which was gradually changing from a volatile transition economy to a more or less stabilized market economy in the EU. Since 2008, the Czech economy has been affected by the global recession and the consequences of the subsequent debt crisis in the euro zone, which have manifested themselves in a repeated increase in volatility of macroeconomic indicators.

Last but not least, all statistics and tests were calculated against the first estimates published by the Czech Statistical Office or Czech National Bank, since it is not possible to estimate the extent of changes in past development through subsequent revisions of time series which cannot usually be divided into components of factual specification of the given ratio and methodological change.

2 FORECAST ERROR MEASUREMENT STATISTICS

The success rate of macroeconomic forecasts is usually evaluated by means of several basic statistics – the average forecasting error, the mean absolute error and Theil's inequality coefficient.[3]

Forecast error (e) or deviation is generally defined as:

$$e_t = F_t - A_t , \tag{1}$$

where F_t is the forecast for the period t and A_t is the real value over time t.

Average forecasting error (AFE) can be regarded as a measure of bias, as it indicates the deviations of forecasts. Positive AFE values indicate systematic or overwhelming overvaluation of forecasts, whereas negative AFE values indicate systematic or overwhelming undervaluation of forecasts. AFE is defined as the average of the forecast errors:

$$AFE = \frac{1}{T}\sum_{t=1}^{T} e_t , \tag{2}$$

with T representing the number of observations.

Mean absolute error (MAE) expresses the average absolute error of the forecast compared to reality. MAE is defined as:

$$MAE = \frac{1}{T}\sum_{t=1}^{T} \left| e_t \right| . \tag{3}$$

---

2 Some analysed indicators have not been included in the Macroeconomic Forecast since the start of publication.

3 Sometimes also the mean percentage error (MPE) and the mean absolute percentage error (MAPE) are used. MPE is defined as an average of the percentage errors and MAPE is defined as an average of the percentage errors. Both statistics ignore the scale of the numbers, however, they can be very unstable and skewed by small values.

Theil's inequality coefficient (TIE) is used for evaluating the success rate of forecasts. The coefficient is defined as the proportion of the mean square variations of analysed forecasts and naïve forecasts, which is used as alternative model (a random walk model):

$$TIE = \frac{\sum_{t=1}^{T}(e_t)^2}{\sum_{t=1}^{T}(A_{t-1} - A_t)^2}. \tag{4}$$

If Theil's coefficient equals 0, the forecast is identical to reality. Value of the coefficient higher than 1 shows that the result of forecasting activities is worse than a naïve forecast. When interpreting the results, it is necessary to take into account the fact that this indicator greatly "penalizes" an isolated considerably worse result compared to the naïve forecast, and conversely, it awards a considerable "bonus" in the event of well-estimated sudden reversals in the development of forecast quantities.

The naïve forecast is a mechanically drawn up forecast where the value of the given indicator for the year of $t + 1$ equals a measured, estimated or forecasted value of this indicator for the year t.

The forecast horizon is understood as the time from publishing the forecast until the end of the forecast period. For any horizons above 15 and up to 24 months, it concerns evaluating an outlook (created by means of extrapolation techniques) whose forecasting information is very limited for understandable reasons.

3 TEST FOR FORECAST ACCURACY

In addition to the basic statistics mentioned above, a statistical test proposed by Diebold and Mariano (1995) for assessing forecast accuracy is used. Diebold-Mariano test compares the forecast accuracy of two forecast methods and it is applicable to non-quadratic loss functions, multi-period forecasts, and forecasts errors that are potentially non-Gaussian, non-zero-mean, serially correlated and contemporaneously correlated.

The asymptotic test introduced by Diebold and Mariano tests the null hypothesis of no difference in the accuracy of two competing forecasts. Suppose two different forecasts y_{1t}, y_{2t}, where t = (1,...n) and let e_{1t}, e_{2t} be the forecast errors of these forecasts. Then the economic loss functions $g(e_{1t})$ and $g(e_{2t})$ are arbitrary functions of the realization and prediction.[4] When denoting a loss differential as $d_t = g(e_{1t}) - g(e_{2t})$, the null hypothesis can be expressed as $H_0 : E(d_t) = 0$. If the expected value of the loss differential is zero, there is no statistical difference between the two forecasts. If the null hypothesis is rejected, the forecast with smaller loss will be chosen.

The Diebold-Mariano test statistic is defined as:

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}}, \tag{5}$$

where $\bar{d} = \frac{\sum_{t=1}^{n} d_t}{n}$ is the sample mean loss differential.

An optimal h-step forecast error will follow a moving average process of order $(h-1)$:

$$e_t = \theta_0\varepsilon_t + \theta_1\varepsilon_{t-1} + \ldots + \theta_{h-1}\varepsilon_{t-h+1}, \tag{6}$$

[4] Some popular economic loss functions are squared error loss $g(e_{it}) = (e_{it})^2$ or absolute error loss $g(e_{it}) = |e_{it}|$, where $i = 1, 2$.

with zero autocovariances for all lags greater than $h-1$. Therefore, the consistent estimate of the asymptotic variance of $\bar{d}$ can be written as:

$$\hat{V}(\bar{d}) = \frac{\left[\hat{\gamma}_0 + 2\sum_{k=1}^{n-1}\hat{\gamma}_k\right]}{n}, \tag{7}$$

where γ_k is an estimate of the k th autocovariance of d_t that can be computed as:

$$\hat{\gamma}_k = \frac{\sum_{t=k+1}^{n}(d_t - \bar{d})(d_{t-k} - \bar{d})}{n} = \frac{\mathrm{cov}(d_t, d_{t-k})}{n}. \tag{8}$$

Under the null hypothesis, DM statistic has an asymptotic standard normal distribution. However, the major drawback of this test statistic is its small sample properties. Simulations showed that DM test statistic is seriously oversized, especially in small samples, so the null hypothesis is being rejected too often. Therefore Harvey, Leybourne and Newbold (1997) proposed modification, which reduces this oversizing:

$$mDM = \sqrt{\frac{n+1-2h+h(h-1)/n}{n}} * DM. \tag{9}$$

The modified DM statistic has a Student's t distribution with $n-1$ degrees of freedom.

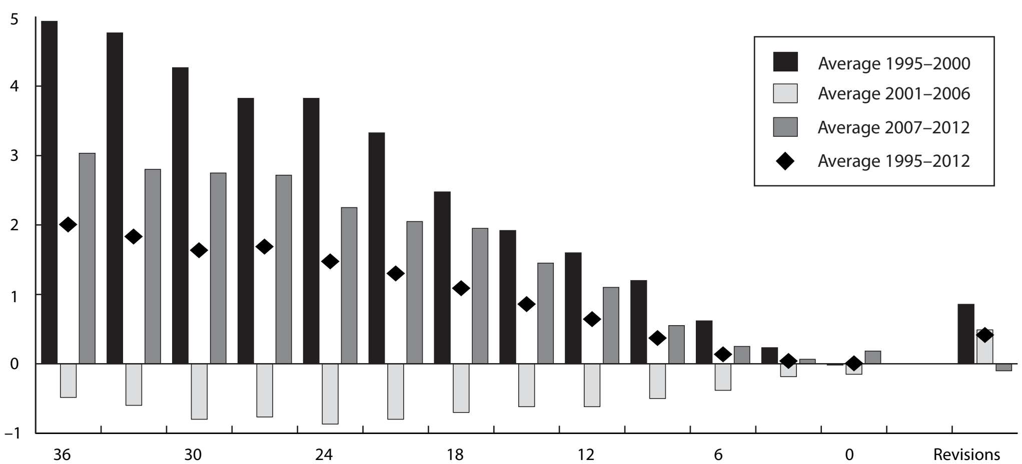
In this analysis, the macroeconomic forecasts are compared with the naïve forecast. Further, commonly used mean squared error loss function is applied and variance is estimated as the long-run variance using a Newey-West method.

4 EVALUATION OF THE MINISTRY OF FINANCE FORECASTS

4.1 Real GDP Growth

In 1995–2000 and 2007–2012 the Ministry of Finance's forecasts overvalued real GDP growth, with forecasts widest of the mark in 1998, 2009 and 2012, when the Czech Republic was in recession. Conversely, in 2001–2006 when the Czech Republic was going through a period of relatively strong and stable economic growth, GDP growth was slightly undervalued, although this undervaluation did not exceed –0.9 p.p.

Figure 1 Average Forecasting Error (in p.p.)



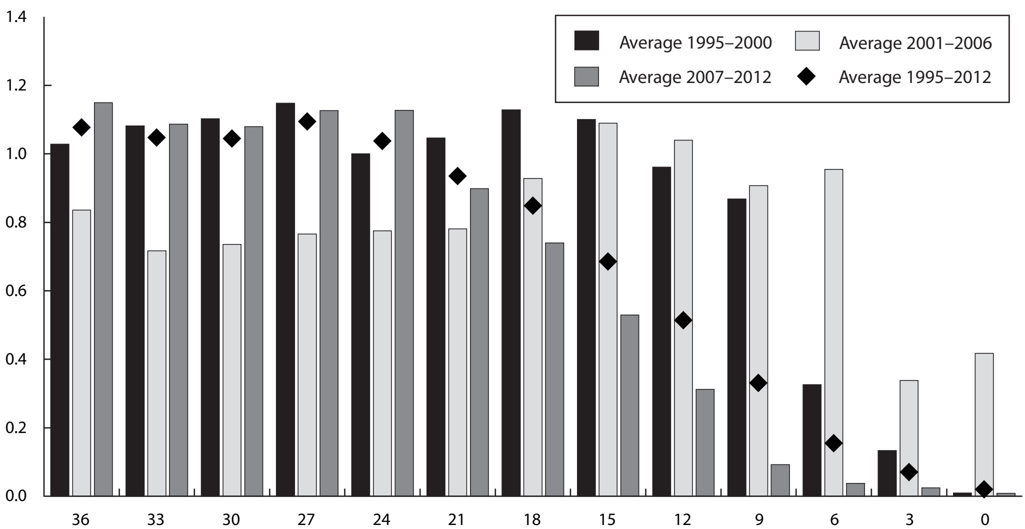
Source: Czech Statistical Office, own calculation

In accordance with results published in the literature and based on the experience of forecasters, it has been proved very difficult, even impossible, to identify the onset of recession in time. In the first and third monitored periods, the mean absolute error exceeded in the horizon over 18 months the limit of 3 p.p., which was caused in particular by the failure to identify recessions in 1998, 2009 and 2012. In the successful period of 2001–2006, the mean absolute error fluctuated below 1.7 p.p. throughout the horizon.

In connection with the so-called great recession at the turn of 2008 and 2009, it is necessary to emphasize, however, that the decline in the domestic economy was caused exclusively by unfavourable development in the external environment. Comparison with the forecasts of other institutions at that time confirms how difficult it was to predict future development.

Theil's coefficient in the forecast horizon beyond 24 months exceeds 1 on average. However, this gradually decreases with a shortening horizon. The analysis proves that the recognisability of future development in an 18-month horizon exceeds only slightly the possibilities of the naïve forecast. It is in this very horizon that the macroeconomic framework of the draft state budget is usually drawn up. This knowledge can also be related to many of the following indicators.

Figure 2 Theil's Coefficient



Source: Czech Statistical Office, own calculation

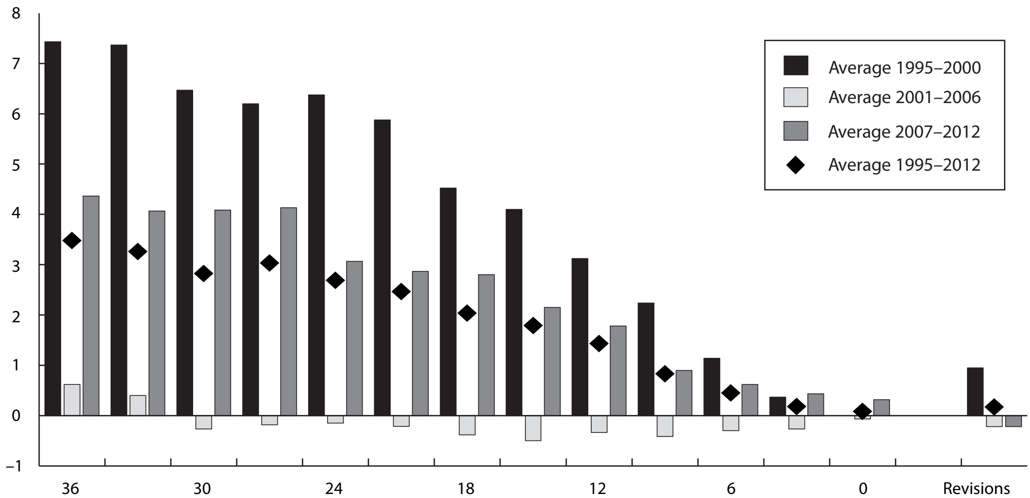
Modified Diebold-Mariano test is even stricter than Theil's coefficient. As can be seen in the Table 8, Modified Diebold-Mariano test showed that there are no the differences between forecast and naïve forecast for 15-month and longer time horizon at 5% level of significance.

4.2 Nominal GDP Growth

From the perspective of the budget process, the most important macroeconomic indicator is nominal GDP. It is used as the denominator of important ratios (e.g. the government sector's balance or debt as a ratio to GDP) and budget revenue forecasts are derived from the size of its components.

As in the case of real GDP growth, nominal GDP growth was overvalued by forecasts in the first and third periods, although the overvaluation in 2007–2012 was likewise considerably lower. Undervaluation of nominal GDP growth in 2001–2006 was only minimal.

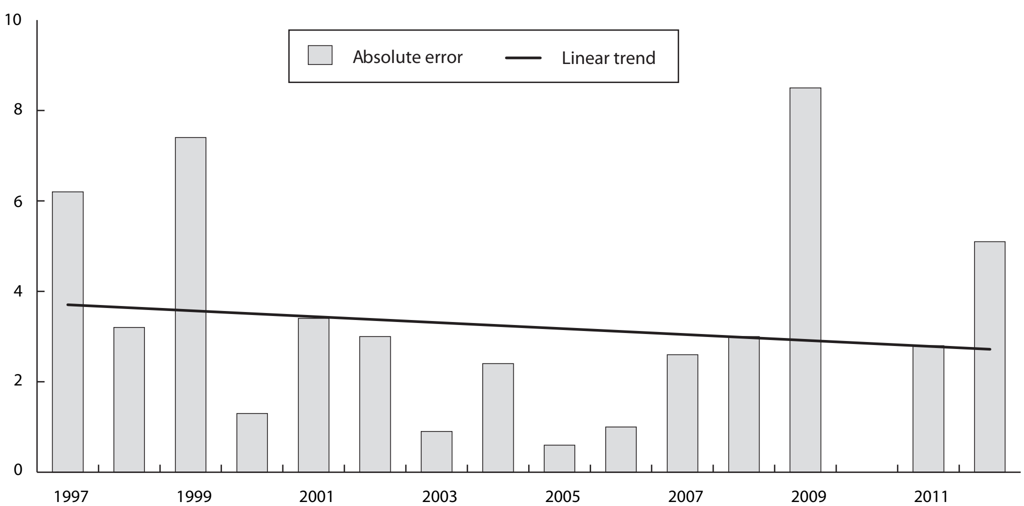
Figure 3 Average Forecasting Error (in p.p.)



Source: Czech Statistical Office, own calculation

In the 18-month horizon representing the starting point for drafting the state budget, the mean absolute error for the whole period reached approximately 3 p.p., although it shows a decreasing tendency during the whole period. Its high values in 1997, 2009 and 2012 were recorded for periods of economic recession, the year 1999 relates to a period of disinflation. The average value of Theil's coefficient in the forecast horizon up to 27 months is lower than 1, while it reaches its lowest values in 2001–2006.

Figure 4 Mean Absolute Error in the 18-Month Horizon (in p.p.)



Source: Czech Statistical Office, own calculation

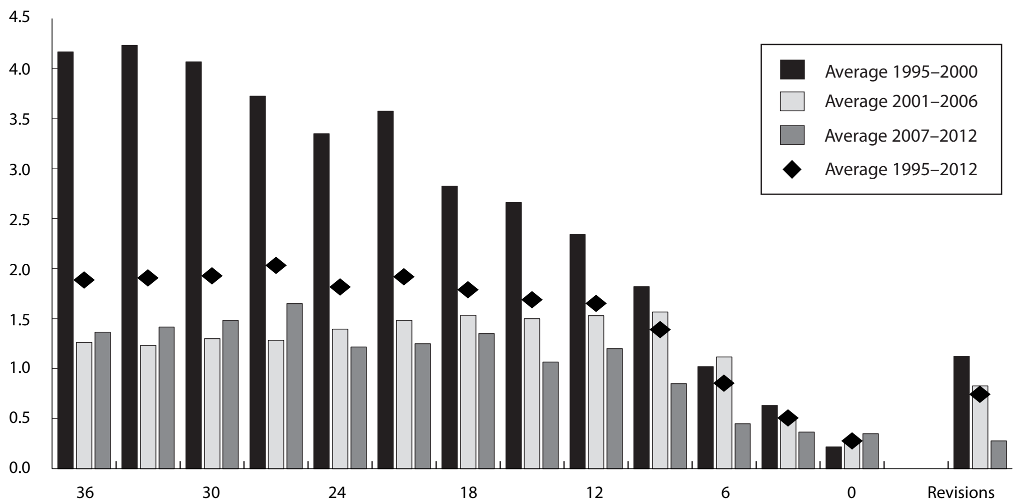
According to the modified Diebold-Mariano test, there are no the differences between forecast and naïve forecast for 15-month and longer time horizon at 5% level of significance.

4.3 GDP Deflator Growth

GDP deflator growth was overvalued in every single monitored period; nevertheless, the average mean error against the actual facts did not exceed 1.4 p.p. throughout the horizon.

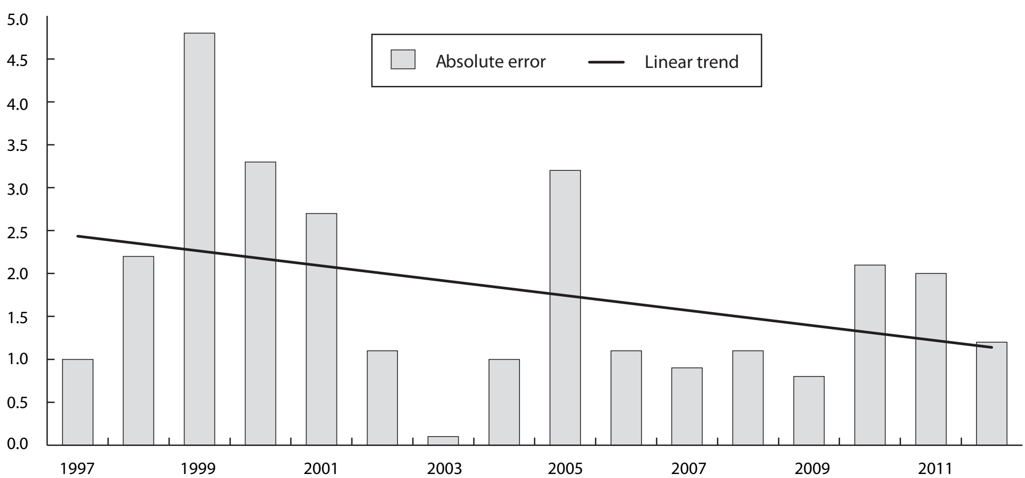
The average mean absolute error did not exceed 2 p.p., and reached its highest values in 1995–2000. The decreasing trend is also confirmed by the graph showing absolute error in the 18-month horizon. The error for 1999 relates to the period of disinflation, when GDP deflator growth decreased from 10.7% in 1998 to 2.4% in 1999. Although a decrease was expected and identified correctly in time, its extent exceeded all expectations.

Figure 5 Mean Absolute Error (in p.p.)



Source: Czech Statistical Office, own calculation

Figure 6 Mean Absolute Error in the 18-Month Horizon (in p.p.)



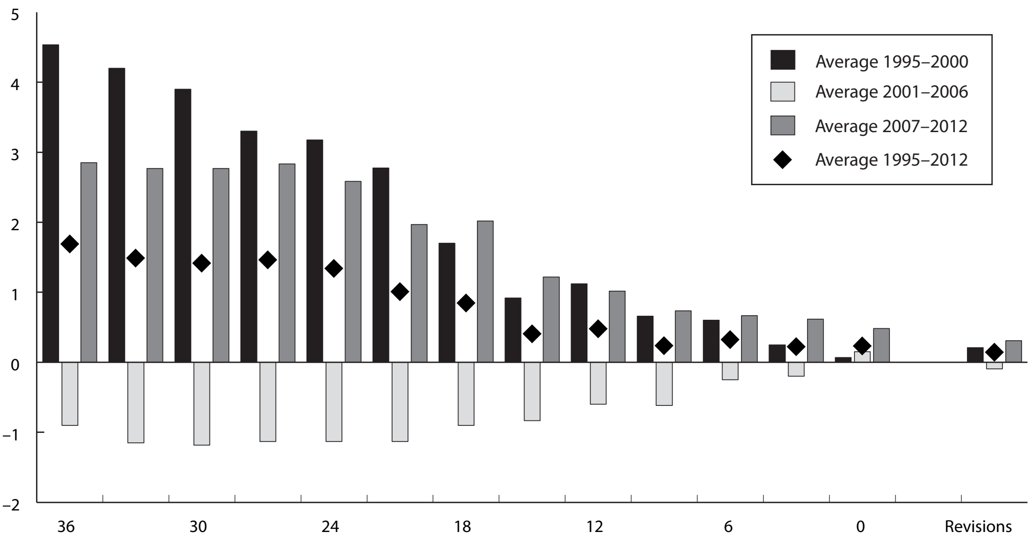
Source: Czech Statistical Office, own calculation

The average Theil's coefficient did not exceed the value of 1.0 throughout the horizon. In the horizon up to 21 months its values decreased gradually in individual periods, thereby highlighting the improvement of forecasts. On the other hand, modified Diebold-Mariano test showed that there are no the differences between forecast and naïve forecast for 18-month and longer time horizon at 5% level of significance, as shown in Table 10.

4.4 Real Private Consumption Growth

While in the first and third monitored periods the growth in household consumption was overvalued, in the second period forecasts were slightly tilted to the downside.

Figure 7 Average Forecasting Error (in p.p.)



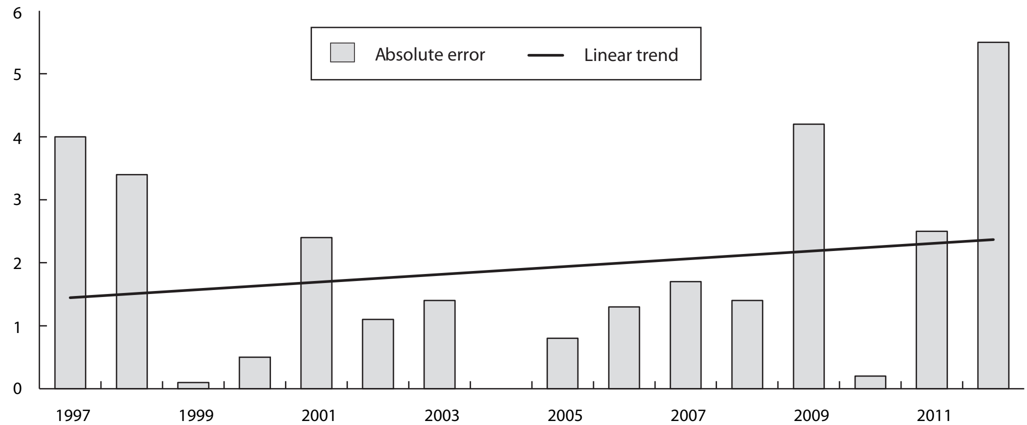
Source: Czech Statistical Office, own calculation

The mean absolute error in individual periods reaches approximately the same values as in case of forecasts of real GDP growth. In the horizon of 2–3 years, it is approximately 3 p.p. on average, whereupon it gradually decreases and drops below 1.5 p.p. within a short period of up to one year.

The absolute error in the 18-month horizon shows an increasing tendency. However, this result is strongly influenced by the imprecise estimate of household consumption in 2012. The extremely low level of consumer confidence in future economic development, together with the implementation of the government's austerity measures, led to cautious behaviour on the part of consumers and to an increase in the rate of savings as a precaution against any further worsening of the economic situation. Thus the decrease in household consumption by 2.1% in 2012 exceeded all expectations. After all, in 2009 during the recession household consumption had even increased by 0.2%!

The average value of Theil's coefficient fluctuated below 1.0 in the horizon up to 18 months. However, in 2007–2012 the coefficient reached considerably higher values than in the other two periods, which was caused in particular by imprecise estimates in 2009 and 2012. According to modified Diebold-Mariano test, there are no the differences between forecast and naïve forecast for 15-month and longer time horizon at 5% level of significance.

Figure 8 Mean Absolute Error in the 18-Month Horizon (in p.p.)



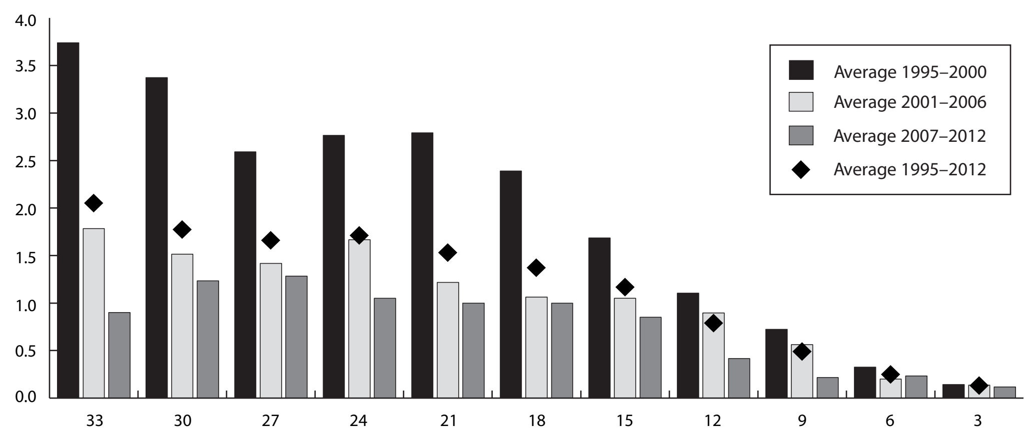
Source: Czech Statistical Office, own calculation

4.5 Average Inflation Rate

Forecasts of inflation in the Macroeconomic Forecast were precise in most cases, since in the horizon up to 30 months the average forecasting error did not exceed 1 p.p. for the whole monitored period. In 1995–2000 and 2001–2006, forecasts slightly overvalued the average inflation rate, while in the second period the overvaluation was higher. Conversely, in 2007–2012 the average mean error achieved negative values, although it did not fall below –0.5 p.p. in any of the horizons.

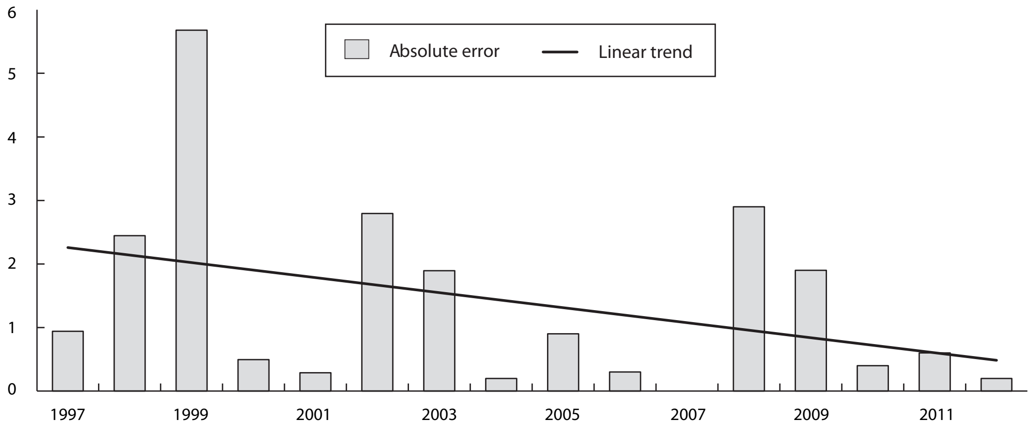
In the horizon up to 30 months, the mean absolute error did not exceed 2 p.p. In the budget horizon of 18 months the mean absolute error has a decreasing tendency. The error for 1999 relates to a period of severe disinflation, when the average inflation rate fell from 10.7% in 1998 to 2.1% in 1999. Although this tendency was identified correctly, its extent exceeded all expectations. The fact that in the budget horizon of 18 months the absolute error did not exceed 1.0 p.p. in 10 out of the 16 monitored years is testimony to the precision of inflation forecasting.

Figure 9 Mean Absolute Error (in p.p.)



Source: Czech Statistical Office, own calculation

Figure 10 Mean Absolute Error in the 18-Month Horizon (in p.p.)



Source: Czech Statistical Office, own calculation

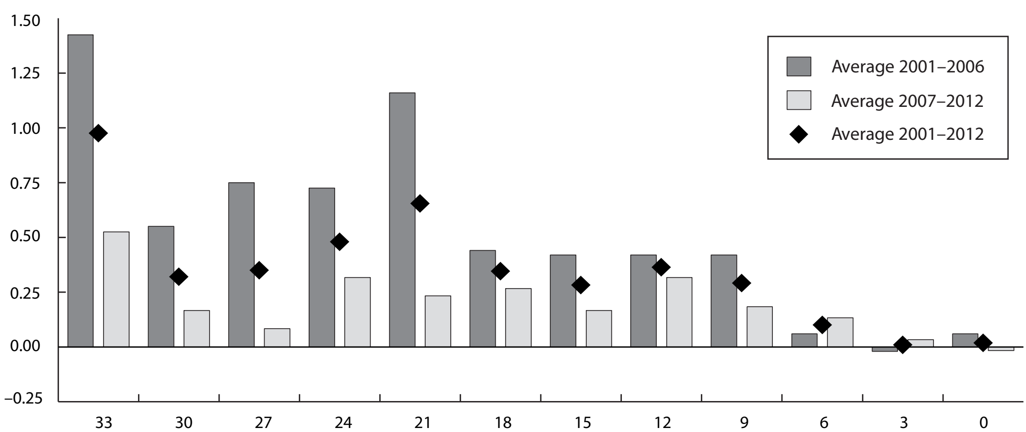
Theil's inequality coefficient for all monitored periods did not exceed 0.85 in the whole time horizon and was 0.15 in the short 1-year period. As can be seen in the Table 12, also modified Diebold-Mariano test showed that there are no the differences between forecast and naïve forecast for 24-month and longer time horizon on 1% level of significance.

4.6 Average Unemployment Rate (LFS)

The unemployment rate according to LFS has only been forecast since 2000, so any comparison of the quality of forecasts over time was possible only for the periods of 2001–2006 and 2007–2012.

The forecasts systematically overvalued the unemployment rate, still the average mean error did not exceed 1.0 p.p. in any time horizon. In 2007–2012, the overvaluation compared to the previous period was considerably lower: the average mean forecasting error did not exceed 0.55 p.p. in any horizon.

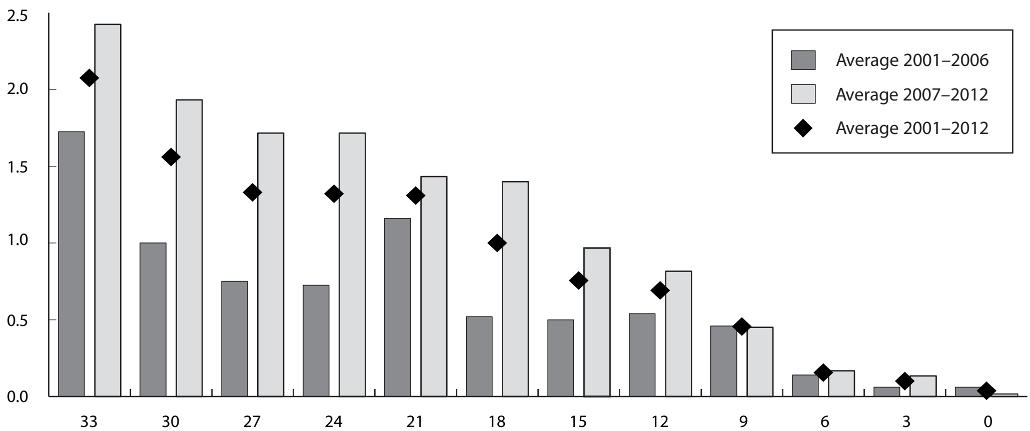
Figure 11 Average Forecasting Error (in p.p.)



Source: Czech Statistical Office, own calculation

The average mean absolute error showed a gradually decreasing tendency. Nonetheless, in 2007–2012 it reached higher values due to the difficultly in forecasting at a time of economic instability compared to the previous period. In the 18-month budget horizon, the mean absolute error has an increasing tendency with respect to imprecise estimates in 2009 and 2007. In 2009, the unemployment rate was undervalued when as a result of the economic recession it increased by 2.3 p.p. compared to the previous year. On the other hand, in 2007 the unemployment rate was overvalued, since strong economic growth resulted in its decrease down to 4.4%. Data for 2004 are missing due to a change in methodology.

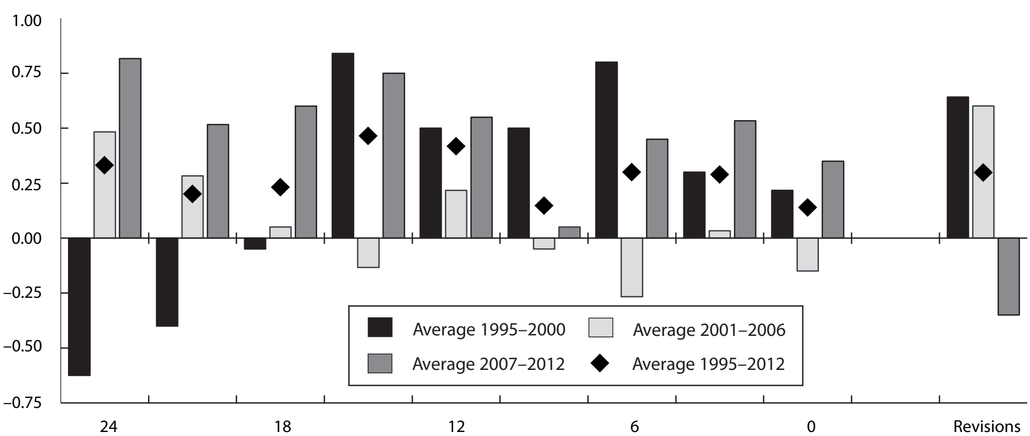
Figure 12 Mean Absolute Error (in p.p.)



Source: Czech Statistical Office, own calculation

These imprecise estimates are also reflected in the higher average value of Theil's coefficient, which exceeds the value of 1.0 in the horizon of 33, 21 and 18 months. Modified Diebold-Mariano test showed that there are no the differences between forecast and naïve forecast for 12-month and longer time horizon at 5% level of significance.

Figure 13 Average Forecasting Error (in p.p.)

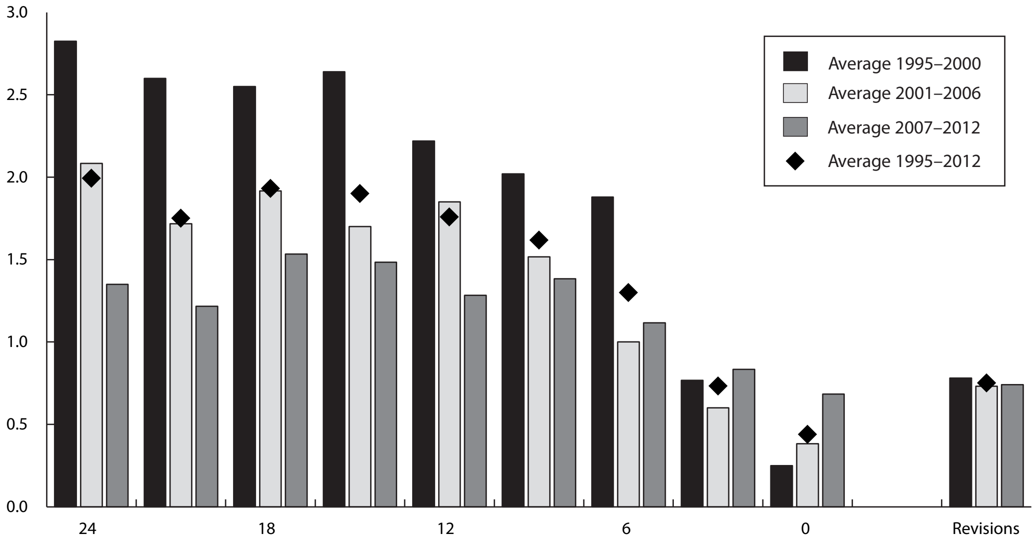


Source: Czech Statistical Office, own calculation

4.7 Current Account Balance to GDP Ratio

During the monitored period, the forecasts overvalued the ratio of the current account balance to GDP. However, the average forecasting error did not exceed 0.5 p.p. on average. The average mean absolute error was between 1 and 2 p.p. in the horizon of 6–24 months, while it was usually the lowest in the third monitored period. Absolute error in the 18-month horizon has a decreasing character.

Figure 14 Mean Absolute Error (in p.p.)



Source: Czech Statistical Office, own calculation

Except for the horizon of 15 months, the average Theil's coefficient was lower than 1. However, it reached its lowest values in the first period, while in 2007–2012 it was even higher than 1 in the horizon of 6–18 months. This phenomenon can largely be attributed to a change in the system of revisions. While revisions were previously on-going, now they occur only once a year. Consequently, the period in which the forecast is based on subsequently revised data is extended.

Modified Diebold-Mariano test showed that there are no the differences between forecast and naïve forecast for 12-month and longer time horizon at 5% level of significance, as is evident from the Table 14.

5 COMPARISON OF RESULTS OF MINISTRY OF FINANCE'S FORECASTS WITH FORECASTS OF INTERNATIONAL INSTITUTIONS

The Ministry of Finance's forecasts were compared with macroeconomic forecasts of the OECD, the European Commission and the International Monetary Fund for 2001–2012 in the horizons corresponding to their mainly half-yearly publishing cycle. The results indicate that the forecast success rate of all institutions does not differ much in essence. The best results are mostly achieved by forecasts from the Ministry of Finance and OECD. The Ministry of Finance's forecasts are the most precise, especially in terms of nominal GDP growth, GDP deflator growth and average inflation rate. On the other hand, the Ministry of Finance's forecasts were the least accurate in the case of unemployment rate.

Table 1 Forecasts of Real GDP Growth (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error				Mean Absolute Error				Theil's Inequality Coefficient			
	MoF	EC	OECD	IMF	MoF	EC	OECD	IMF	MoF	EC	OECD	IMF
27 months	**0.98**	1.13	1.18	-	**2.49**	2.57	2.62	-	1.06	**0.99**	1.11	-
21 months	**0.63**	0.95	1.05	0.69	**2.34**	2.47	2.44	2.45	0.88	0.93	**0.83**	0.89
15 months	**0.42**	0.55	0.61	0.53	2.00	2.05	**1.79**	2.16	0.57	0.56	**0.45**	0.62
9 months	**0.03**	–0.03	–0.10	–0.26	1.09	1.03	**0.75**	0.99	0.15	0.14	**0.08**	0.12
3 months	–0.06	–0.17	**–0.02**	–0.28	0.51	**0.43**	0.47	0.63	0.04	0.04	**0.04**	0.07

Note: The best estimate is marked in bold.
Source: Czech Statistical Office, European Commission, OECD, IMF, own calculation

Table 2 Forecasts of Nominal GDP Growth (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error			Mean Absolute Error			Theil's Inequality Coefficient		
	MoF	EC	OECD	MoF	EC	OECD	MoF	EC	OECD
27 months	**1.98**	2.49	2.09	3.36	3.64	**3.17**	1.18	1.08	**0.99**
21 months	**1.33**	2.05	2.20	**2.76**	2.94	2.82	0.85	1.03	**0.67**
15 months	**0.83**	1.36	1.58	**2.53**	2.67	2.53	**0.60**	0.63	0.71
9 months	**0.24**	0.36	0.91	1.78	**1.77**	1.96	**0.32**	0.41	0.51
3 months	**0.08**	0.14	0.11	**0.67**	1.39	0.78	**0.06**	0.29	0.08

Note: The best estimate is marked in bold.
Source: Czech Statistical Office, European Commission, OECD, own calculation

Table 3 Forecasts of GDP Deflator Growth (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error			Mean Absolute Error			Theil's Inequality Coefficient		
	MoF	EC	OECD	MoF	EC	OECD	MoF	EC	OECD
27 months	0.93	1.13	**0.82**	1.47	1.45	**1.02**	1.56	0.97	**0.84**
21 months	**0.67**	1.03	1.09	1.37	1.43	**1.15**	0.56	0.78	**0.33**
15 months	**0.35**	0.86	0.90	**1.28**	1.39	1.32	**0.40**	0.65	0.55
9 months	**0.21**	0.50	0.98	**1.21**	1.32	1.53	**0.33**	0.63	0.66
3 months	0.11	0.32	**0.11**	**0.44**	1.14	0.51	**0.05**	0.44	0.06

Note: The best estimate is marked in bold.
Source: Czech Statistical Office, European Commission, OECD, own calculation

Table 4 Forecasts of Private Consumption Growth (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error			Mean Absolute Error			Theil's Inequality Coefficient		
	MoF	EC	OECD	MoF	EC	OECD	MoF	EC	OECD
27 months	**0.85**	2.19	1.51	2.52	2.81	**2.37**	1.32	1.37	**1.27**
21 months	**0.42**	1.45	0.93	**2.05**	2.33	2.05	**1.28**	1.45	1.50
15 months	**0.19**	1.11	0.50	1.76	1.93	**1.75**	0.81	0.91	**0.73**
9 months	**0.06**	0.39	–0.13	1.19	1.21	**0.94**	0.50	0.48	**0.29**
3 months	**0.21**	0.32	0.30	**0.61**	0.62	0.75	0.11	**0.11**	0.13

Note: The best estimate is marked in bold.
Source: Czech Statistical Office, European Commission, OECD, own calculation

Table 5 Forecasts of Average Inflation Rate (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error			Mean Absolute Error			Theil's Inequality Coefficient		
	MoF	OECD	IMF	MoF	OECD	IMF	MoF	OECD	IMF
27 months	0.52	**0.38**	-	1.35	**1.35**	-	0.78	**0.78**	-
21 months	**0.41**	0.51	0.53	**1.11**	1.30	1.38	**0.48**	0.51	0.62
15 months	**0.47**	0.53	0.54	0.95	**0.94**	1.20	0.33	**0.29**	0.40
9 months	**0.07**	0.45	0.37	**0.39**	0.59	0.51	**0.06**	0.11	0.11
3 months	**0.02**	0.12	0.17	**0.13**	0.19	0.33	**0.01**	0.01	0.03

Note: The best estimate is marked in bold.
Source: Czech Statistical Office, OECD, IMF, own calculation

Table 6 Forecasts of Average Unemployment Rate LFS (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error			Mean Absolute Error			Theil's Inequality Coefficient		
	MoF	EC	OECD	MoF	EC	OECD	MoF	EC	OECD
27 months	0.35	0.28	**0.23**	1.33	1.30	**1.26**	0.90	0.89	**0.81**
21 months	0.65	**0.49**	0.67	1.31	**1.21**	1.27	1.21	**0.83**	1.04
15 months	0.28	**0.28**	0.28	0.75	**0.70**	0.80	0.76	0.71	**0.62**
9 months	**0.29**	0.31	0.42	0.45	0.47	**0.44**	0.31	**0.31**	0.35
3 months	**0.01**	0.18	0.07	**0.10**	0.18	0.15	**0.02**	0.08	0.03

Note: The best estimate is marked in bold.
Source: Czech Statistical Office, European Commission, OECD, own calculation

Table 7 Forecasts of Current Account Balance to GDP Ratio (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error			Mean Absolute Error			Theil's Inequality Coefficient		
	MoF	OECD	IMF	MoF	OECD	IMF	MoF	OECD	IMF
27 months	3.70	**0.25**	-	3.70	**1.63**	-	2.75	**0.91**	-
21 months	0.40	0.55	**–0.06**	1.47	1.65	**1.01**	0.86	1.41	**0.76**
15 months	0.31	0.36	**0.23**	1.59	1.89	**1.26**	1.32	1.48	**1.09**
9 months	**0.00**	0.48	0.04	1.45	1.33	**1.08**	1.18	1.09	**0.67**
3 months	0.28	0.22	**0.19**	**0.72**	1.02	0.93	**0.35**	0.62	0.59

Note: The best estimate is marked in bold.
Source: Czech Statistical Office, OECD, IMF, own calculation

As far as consumer prices are concerned, the EC forecasts HICP, which cannot be compared with the national CPI. In the forecasts of the EC, current account balance to GDP ratio is defined in national accounts terms. The IMF forecasts include only the forecasts for real GDP growth, inflation rate and the current account balance to GDP ratio.

CONCLUSION

Based on the forecast error measurement statistics, it is possible to say that for most macroeconomic indicators forecasts contain valid data in a horizon of approximately up to 18 months (it is important to note that the macroeconomic framework of the draft state budget is usually drawn up in this horizon). In longer horizons, however, the objective is geared more towards determining the expected trends of economic development. The results of the modified Diebold-Mariano test are even stricter. According to the results, the most macroeconomic indicators forecasts contain valid data only in a horizon of approximately up to 12 months at 5% level of significance. In this case, it can be generalized that modified Diebold-Mariano test confirms null hypothesis of no difference in the accuracy of Ministry of Finance's Macroeconomic Forecasts and naïve forecast at 5% level of significance for most macroeconomic indicators at 0.6 to 0.8 value of Theil's coefficient.

As far as the development of forecast precision over time is concerned, it is apparent that forecast precision increased in the second and third monitored periods (2001–2006, 2007–2012) compared to the first period (1995–2000). In this context, however, it must be pointed out that forecasting future economic development is considerably more difficult at a time of economic crisis and recession than in a period of stable economic growth. This fact was the main reason for several imprecise forecasts in 2007–2012.

Assessment of the history of the Ministry of Finance's Macroeconomic Forecasts also has showed that they are fully comparable to the forecasts of renowned international institutions, and in a number of cases even surpass them. The Ministry of Finance usually publishes its forecasts earlier than the other institutions included in this comparison.

References

BATCHELOR, R. *The IMF and OECD versus Consensus Forecasts* [online]. London: City University Business School, 2000. [cit. 1. 7. 2013]. <http://www.consensuseconomics.com/Batchelor_Study.pdf>.

CARSTENSEN, K., WOHLRABE, K., ZIEGLER, C. *Predictive Ability of Business Cycle Indicators under Test: A Case Study for the Euro Area Industrial Production* [online]. CESifo Working Paper No. 3158, Institute for Economic Research at the University of Munich, August 2010. [cit. 17. 12. 2013]. <http://epp.eurostat.ec.europa.eu/portal/pls/portal/!PORTAL.wwpob_page.show?_docname=2304447.PDF>.

DIEBOLD, F. X., MARIANO, R. S. Comparing Predictive Accuracy [online]. *Journal of Business and Economic Statistics,* 1995, Vol. 13, No. 3, pp. 253–263. [cit. 1. 7. 2013]. <http://www.est.uc3m.es/esp/nueva_docencia/comp_col_get/lade/tecnicas_prediccion/Practicas0708/Comparing%20Predictive%20Accuracy%20(Dielbold).pdf>.

GARCIA-FERRER, A., DE JUAN, A., PONCELA, P., BUJOSA, M. Monthly forecasts of integrated public transport systems: The case of the Madrid Metropolitan Area: The Case of the Madrid Metropolitan Area. *Journal of Transportation and Statistics,* 2004, Vol. 7, No. 1, pp. 39–59.

HARVEY, D. *The Evaluation of Economic Forecasts* [online]. Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy, October 1997. [cit. 18. 12. 2013]. <http://etheses.nottingham.ac.uk/198/1/David_Harvey_PhD_Thesis.pdf>.

HARVEY, D., LEYBOURNE, S., NEWBOLD, P. Testing the Equality of Prediction Mean Squared Errors. *International Journal of Forecasting,* June 1997, Vol. 13, Issue 2, pp. 281–291.

LUGER, R. *Exact Tests of Equal Forecast Accuracy with an Application to the Term Structure of Interest Rates* [online]. Working paper, Bank of Canada, Ottawa, 2004. [cit. 21. 11. 2013]. <http://www.bankofcanada.ca/wp-content/uploads/2010/02/wp04-2.pdf>.

MARIANO, R. S. *Testing Forecast Accuracy* [online]. University of Pennsylvania, 2000. [cit. 1. 7. 2013]. <http://projects.chass.utoronto.ca/link/200010/papers/testforecast.pdf>.

MELANDER, A., SISMANIDIS, G., GRENOUILLEAU, D. *The Track Record of the Commission's Forecasts – an Update* [online]. Brussels: European Commission, 2007. [cit. 1. 7. 2013]. <http://ec.europa.eu/economy_finance/publications/publication9291_en.pdf>.

NOVOTNÝ, F., RAKOVÁ, M. *Assessment of Consensus Forecasts Accuracy: The Czech National Bank Perspective* [online]. Prague: Czech National Bank, 2010. [cit. 1. 7. 2013]. <https://www.cnb.cz/miranda2/export/sites/www.cnb.cz/en/research/research_publications/cnb_wp/download/cnbwp_2010_14.pdf>.

OSTERLOH, S. *Accuracy and Properties of German Business Cycle Forecasts* [online]. Mannheim, ZEW Discussion Paper No. 06–087, 2006. [cit. 1. 7. 2013]. <http://econstor.eu/bitstream/10419/24543/1/dp06087.pdf>.

ANNEX

Table 8 Forecasts of Real GDP Growth (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error				Mean Absolute Error				TIE	DM test
	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2012
36 months	2.01	4.93	–0.48	3.03	2.98	4.93	1.15	3.83	1.08	0.12
33 months	1.83	4.77	–0.60	2.80	2.87	4.77	1.23	3.57	1.05	0.10
30 months	1.63	4.27	–0.80	2.75	2.90	4.27	1.53	3.58	1.04	0.12
27 months	1.69	3.83	–0.77	2.72	2.88	4.03	1.50	3.48	1.09	0.33
24 months	1.48	3.83	–0.87	2.25	2.70	3.98	1.47	3.08	1.04	0.15
21 months	1.30	3.33	–0.80	2.05	2.75	3.98	1.63	3.05	0.93	–0.30
18 months	1.09	2.48	–0.70	1.95	2.63	3.53	1.53	3.12	0.85	–0.77
15 months	0.86	1.92	–0.62	1.45	2.18	2.60	1.35	2.65	0.69	–1.70
12 months	0.64	1.60	–0.62	1.10	1.77	2.24	1.22	1.93	0.51	–2.55**
9 months	0.37	1.20	–0.50	0.55	1.38	2.08	0.97	1.22	0.33	–3.54***
6 months	0.14	0.62	–0.38	0.25	0.90	1.26	0.75	0.75	0.15	–4.58***
3 months	0.04	0.23	–0.18	0.07	0.59	0.77	0.45	0.57	0.07	–5.31***
0 month	0.01	–0.02	–0.15	0.18	0.33	0.28	0.38	0.32	0.02	–5.89***
Revisions	0.40	0.86	0.49	–0.14	0.79	1.47	0.66	0.25	x	x

Note: Stars indicate if the null hypothesis of the same forecast accuracy of the compared forecasts can be rejected at these level of significance: *** 1%, ** 5%, * 10%.
Source: Czech Statistical Office, own calculation

Table 9 Forecasts of Nominal GDP Growth (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error				Mean Absolute Error				TIE	DM test
	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2012
36 months	3.48	7.43	0.62	4.37	4.03	7.43	1.02	5.33	1.04	0.09
33 months	3.26	7.37	0.40	4.07	3.97	7.37	1.20	5.03	1.06	0.16
30 months	2.82	6.47	–0.27	4.08	3.94	6.47	1.50	5.12	1.04	0.13
27 months	3.03	6.20	–0.18	4.13	4.07	6.20	1.55	5.17	0.96	–0.21
24 months	2.69	6.38	–0.15	3.07	3.71	6.38	1.58	4.07	0.97	–0.14
21 months	2.46	5.88	–0.22	2.87	3.54	5.88	1.78	3.73	0.91	–0.46
18 months	2.04	4.53	–0.38	2.80	3.21	4.53	1.88	3.67	0.86	–0.75
15 months	1.79	4.10	–0.50	2.15	2.99	4.10	1.87	3.18	0.75	–1.47
12 months	1.43	3.12	–0.33	1.78	2.49	3.12	1.83	2.62	0.59	–2.17**
9 months	0.83	2.24	–0.42	0.90	1.94	2.32	1.98	1.57	0.36	–3.45***
6 months	0.45	1.14	–0.30	0.62	1.13	1.22	1.27	0.92	0.15	–4.76***
3 months	0.18	0.37	–0.27	0.43	0.83	1.17	0.50	0.83	0.07	–5.96***
0 month	0.08	0.00	–0.07	0.32	0.36	0.33	0.30	0.45	0.01	–6.66***
Revisions	0.15	0.95	–0.22	–0.29	0.87	1.45	0.82	0.34	x	x

Note: Stars indicate if the null hypothesis of the same forecast accuracy of the compared forecasts can be rejected at these level of significance: *** 1%, ** 5%, * 10%.
Source: Czech Statistical Office, own calculation

Table 10 Forecasts of GDP Deflator Growth (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error				Mean Absolute Error				TIE	DM test
	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2012
36 months	1.32	2.03	1.05	1.22	1.88	4.17	1.26	1.37	0.85	–0.13
33 months	1.25	2.10	0.93	1.15	1.91	4.23	1.23	1.42	0.92	–0.09
30 months	1.06	1.80	0.53	1.22	1.93	4.07	1.30	1.48	0.83	–0.26
27 months	1.19	1.98	0.55	1.32	2.03	3.73	1.28	1.65	0.81	–0.44
24 months	1.14	2.15	0.73	0.88	1.82	3.35	1.40	1.22	0.77	–0.53
21 months	1.04	2.18	0.58	0.75	1.92	3.58	1.48	1.25	0.71	–0.91
18 months	0.84	1.73	0.30	0.78	1.79	2.83	1.53	1.35	0.57	–1.47
15 months	0.81	1.90	0.10	0.60	1.69	2.66	1.50	1.07	0.45	–2.59**
12 months	0.69	1.26	0.30	0.60	1.65	2.34	1.53	1.20	0.36	–3.71***
9 months	0.40	0.86	0.07	0.35	1.39	1.82	1.57	0.85	0.26	–4.54***
6 months	0.28	0.42	0.08	0.35	0.85	1.02	1.12	0.45	0.11	–5.72***
3 months	0.09	0.07	–0.12	0.33	0.51	0.63	0.52	0.37	0.03	–6.74***
0 month	0.05	–0.02	0.04	0.12	0.28	0.22	0.26	0.35	0.01	–7.62***
Revisions	–0.29	0.00	–0.70	–0.16	0.73	1.12	0.83	0.25	x	x

Note: Stars indicate if the null hypothesis of the same forecast accuracy of the compared forecasts can be rejected at these level of significance: *** 1%, ** 5%, * 10%.
Source: Czech Statistical Office, own calculation

Table 11 Forecasts of Private Consumption Growth (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error				Mean Absolute Error				TIE	DM test
	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2012
36 months	1.69	4.53	–0.90	2.85	2.79	4.53	1.20	3.52	1.01	0.01
33 months	1.49	4.20	–1.15	2.77	2.79	4.20	1.45	3.43	1.09	0.20
30 months	1.41	3.90	–1.18	2.77	2.81	3.90	1.65	3.43	1.11	0.3
27 months	1.46	3.30	–1.13	2.83	2.71	3.30	1.53	3.50	1.21	0.69
24 months	1.34	3.18	–1.13	2.58	2.54	3.18	1.43	3.22	1.17	0.61
21 months	1.01	2.78	–1.13	1.97	2.23	2.78	1.47	2.63	1.11	0.39
18 months	0.84	1.70	–0.90	2.02	1.91	2.00	1.17	2.58	0.88	–0.45
15 months	0.41	0.92	–0.83	1.22	1.75	1.72	1.23	2.28	0.58	–2.01*
12 months	0.48	1.12	–0.60	1.02	1.36	1.28	0.97	1.82	0.46	–2.42**
9 months	0.24	0.66	–0.62	0.73	1.15	1.06	1.05	1.33	0.35	–2.81**
6 months	0.32	0.60	–0.25	0.67	0.81	0.72	0.68	1.00	0.18	–3.18***
3 months	0.22	0.25	–0.20	0.62	0.64	0.72	0.57	0.65	0.11	–4.86***
0 month	0.23	0.07	0.15	0.48	0.42	0.40	0.38	0.48	0.05	–5.75***
Revisions	0.11	0.21	–0.09	0.23	0.72	0.96	0.51	0.69	x	x

Note: Stars indicate if the null hypothesis of the same forecast accuracy of the compared forecasts can be rejected at these level of significance: *** 1%, ** 5%, * 10%.
Source: Czech Statistical Office, own calculation

Table 12 Forecasts of Average Inflation Rate (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error				Mean Absolute Error				TIE	DM test
	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2012
33 months	1.09	1.11	1.78	–0.30	2.05	3.74	1.78	0.90	0.83	0.00
30 months	0.56	0.81	1.38	–0.40	1.77	3.37	1.51	1.23	0.68	–0.76
27 months	0.52	0.55	1.42	–0.38	1.66	2.59	1.42	1.28	0.60	–1.33
24 months	0.77	1.22	1.67	–0.42	1.71	2.77	1.67	1.05	0.64	–1.42
21 months	0.59	1.15	1.12	–0.30	1.53	2.79	1.22	1.00	0.46	–3.16***
18 months	0.44	0.70	0.80	–0.10	1.37	2.39	1.06	1.00	0.40	–3.92***
15 months	0.54	0.73	0.98	–0.05	1.17	1.68	1.05	0.85	0.37	–4.54***
12 months	0.37	0.39	0.73	–0.02	0.79	1.10	0.90	0.42	0.14	–6.16***
9 months	0.09	0.13	0.27	–0.12	0.49	0.72	0.56	0.22	0.05	–7.31***
6 months	0.03	–0.07	0.17	–0.03	0.25	0.33	0.20	0.23	0.01	–8.19***
3 months	0.04	0.06	0.13	–0.08	0.13	0.14	0.13	0.12	0.00	–8.66***

Note: Stars indicate if the null hypothesis of the same forecast accuracy of the compared forecasts can be rejected at these level of significance: *** 1%, ** 5%, * 10%.
Source: Czech Statistical Office, own calculation

Table 13 Forecasts of Average Unemployment Rate LFS (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error			Mean Absolute Error			TIE	DM test
	2001–2012	2001–2006	2007–2012	1995–2012	2001–2006	2007–2012	2001–2012	2001–2012
33 months	0.98	1.43	0.53	2.08	1.73	2.43	1.36	x
30 months	0.32	0.55	0.17	1.56	1.00	1.93	0.98	x
27 months	0.35	0.75	0.08	1.33	0.75	1.72	0.90	x
24 months	0.48	0.73	0.32	1.32	0.73	1.72	0.85	–0.24
21 months	0.65	1.16	0.23	1.31	1.16	1.43	1.21	0.61
18 months	0.35	0.44	0.27	1.00	0.52	1.40	1.06	0.15
15 months	0.28	0.42	0.17	0.75	0.50	0.97	0.76	–0.66
12 months	0.36	0.42	0.32	0.69	0.54	0.82	0.68	–0.97
9 months	0.29	0.42	0.18	0.45	0.46	0.45	0.31	–2.30**
6 months	0.10	0.06	0.13	0.15	0.14	0.17	0.05	–3.59***
3 months	0.01	–0.02	0.03	0.10	0.06	0.13	0.02	–4.04***
0 month	0.02	0.06	–0.02	0.04	0.06	0.02	0.01	–4.51***

Note: Stars indicate if the null hypothesis of the same forecast accuracy of the compared forecasts can be rejected at these level of significance: *** 1%, ** 5%, * 10%.
Source: Czech Statistical Office, own calculation

Table 14 Forecasts of Current Account Balance to GDP Ratio (average forecasting error and mean absolute error in p.p.)

	Average Forecasting Error				Mean Absolute Error				TIE	DM test
	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2000	2001–2006	2007–2012	1995–2012	1995–2012
24 months	0.33	–0.63	0.48	0.82	1.99	2.83	2.08	1.35	0.85	–0.86
21 months	0.20	–0.40	0.28	0.52	1.75	2.60	1.72	1.22	0.81	–1.03
18 months	0.23	–0.05	0.05	0.60	1.93	2.55	1.92	1.53	0.91	–0.49
15 months	0.46	0.84	–0.13	0.75	1.90	2.64	1.70	1.48	1.04	0.25
12 months	0.42	0.50	0.22	0.55	1.76	2.22	1.85	1.28	0.86	–0.95
9 months	0.15	0.50	–0.05	0.05	1.62	2.02	1.52	1.38	0.74	–1.76**
6 months	0.30	0.80	–0.27	0.45	1.30	1.88	1.00	1.12	0.55	–3.05***
3 months	0.29	0.30	0.03	0.53	0.73	0.77	0.60	0.83	0.18	–6.88***
0 month	0.14	0.22	–0.15	0.35	0.44	0.25	0.38	0.68	0.05	–9.40***
Revisions	0.30	0.64	0.60	–0.35	0.75	0.78	0.73	0.74	x	x

Note: Stars indicate if the null hypothesis of the same forecast accuracy of the compared forecasts can be rejected at these level of significance: *** 1%, ** 5%, * 10%.
Source: Czech Statistical Office, own calculation

Relation between Composite Indicators and Estimates of Quarterly GDP Changes: Case of the Czech Republic

Jan Zeman[1] | *University of Economics, Prague, Czech Republic*

Abstract

Gross Domestic Product (GDP) represents the basic indicator of macroeconomic performance of the Czech economy and its importance is growing. The need to get the information on its development as quickly as possible for the necessary government acts is unquestionable. Nevertheless, the time taken to publish its first quarterly estimate of growth rate is significantly longer (45 days after the reference quarter) in comparison to some other countries such as the USA and the United Kingdom.
The aim of this paper is to assess the relationship between composite leading indicator (CLI), composite coincidence indicator (CCI) and the development of GDP followed by verification of a predictive ability of these composite indicators. The relationships between GDP and indicators available in this 30-day period which could enter to this CLI and CCI are analysed by the advanced methods of time series analysis.

Keywords

Gross Domestic Product, composite indicator, business tendency surveys, co-integration analysis

JEL code

C43, C83, O47

INTRODUCTION

Preliminary estimates of quarterly Gross Domestic Product (GDP) are designed in many countries to meet the growing pressure on the fastest economic data. These estimates are usually based on incomplete data and various modelling techniques. It is necessary to find a compromise between the two most important requirements – timeliness and quality.

Quarterly GDP is part of the quarterly national accounts, which represent an interconnected system of data on transactions, accounts and balancing items collected on a quarterly basis. In the Czech Republic, these quarterly accounts as well as annual ones are published by the Czech Statistical Office. Regarding terminology, there are some terms used worldwide that the user may not clearly understand at a first glance. Such terms include so-called "flash estimates" which could be compared to our Czech

1 University of Economics, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. E-mail: jan.zeman@vse.cz, phone: (+420)224095435.

preliminary estimates mentioned above. Eurostat (2003) defines a *flash estimate* as the "earliest picture of the economy with regard to the concepts of national accounts published as soon as possible after the end of the quarter". Another term that may confuse users is the *preliminary estimate* used by the Czech Statistical Office. However, its definition and publishing as the first estimate corresponds more to the term *flash estimate*.

Many authors have tried to construct an indicator that could predict the development of GDP in the near future with a certain amount of accuracy and quality and this paper deals with this idea as well. The aim of the paper is to construct and verify CLI that should anticipate and predict, how GDP will develop in the near future and CCI that should develop consistently with the economic cycle and that can be composed of economic indicators with data available in a period not exceeding 30 days after the end of the quarter.

The paper is organized as follows. The first section offers a summary of current knowledge on this issue, then the data and methods used for its evaluation and the course of the analysis itself are described. In the third section of this paper, key results of the analysis are presented with subsequent verification regarding the actual development of the business cycle that constitutes the last section. Finally, additional procedures linked to this analysis are suggested.

1 STATE OF THE ART

There are four main papers in the current research focused on preliminary estimates of GDP in the Czech Republic. The first attempt to construct GDP estimate was made by Jílek and Vojta (2001). They attempted to construct estimates of GDP at chain-linked prices (as aggregated GDP) without explicitly expressed structure by production or expenditure estimation method. The structure itself contains an algorithm calculating the estimate. The analysis is based on seasonally unadjusted estimate and GDP development is estimated in relation to the same period of the previous year. The algorithm consists of selection of monthly sales indicators for sectors most closely matching the profile manufacturing and reducible to fixed prices. As the next step annual indices of quarterly sales are calculated followed by the gross value added to the Sales ratio in previous years. The last step includes the calculation of shares of each sector on GDP at basic prices in the same quarter of the previous year and these shares represent weights used for summarization of the results for each sector.

The name *user signal estimate* used by Jilek and Vojta has its origin in the fact that this estimate can be realized by anybody with using the data publicly available and usually published and it does not use any additional information from the Czech Statistical Office.

This estimate has been improved by so-called *Improved User Signal Estimate* by the same authors (Jilek and Vojta, 2003).

In this case authors build on their previous work from 2001 and construct an improved estimate. The need of such estimate is justified by experimental calculations which results indicate high variances of signal estimates from current estimates of gross value added in individual sectors.

The authors decided to construct a global signal estimate and not to calculate individual industrial gross value added estimates. The global signal estimate is based on the total sales index calculated as a weighted average of industrial indices of sales, while the weights are represented by the shares of sectors in the gross value added. Unlike previous paper, the authors decided to assess the relationship of changes in sales and changes in GDP by decreasing scale constructed using simple regression relationship.

Both above papers permit to construct an estimate of roughly 50-day delay after the reference quarter. The obvious question is whether it would be possible to construct a preliminary estimate of GDP even earlier (e.g. about 30 days after the reference quarter).

According to this requirement, Jan Fischer (Jan Fischer et al., 2002) and his team contributed by analysis of the relationship among the business cycle balances and gross value added. The analysis deals with an initial thesis stating that there is not enough information about production for construction the estimate until 30 days after the reference quarter. Regarding the international practice, it is usual to use a set of business cycle expectations.

The authors have compiled regression equations with industrial gross value added in manufacturing and in construction, respectively as a response variable and chosen combinations of business cycle expectations balances series as explanatory variables. The paper offers an important finding that the coefficient of determination is not a suitable indicator for assessing the quality of predictions. The essential issue is the low quality of pseudo-predictions.

The last existing attempt to construct quarterly preliminary estimates of GDP was made by Jakub Fischer (2005). This methodology regards the character and information capability of the official estimates and regression analysis is used.

Author used gross value added at basic chain-linked prices of 1995 as a response variable and chose 10 explanatory variables. Its list can be found in Fischer (2005).

Models and their suitability were assessed by construction of pseudo-predictions while all series have been reduced by the value of the last known quarter and after estimation of regression parameters the estimate for the last known quarter was constructed. All these estimates were confronted with the official 70-day estimate. Author chose pseudo-predictions' absolute deviation from official estimate as appropriateness criterion of the model. The best preliminary estimate was based on series of the lagged response variable, rail freight series and indicator of confidence in trade.

The main conclusion is the fact that it is not advisable to use only results of business tendency surveys for construction of the estimate and it is not appropriate to work with five-year and longer time series.

Regarding the issue of estimating the GDP changes utilizing composite indicators, there are several documents suggesting alternative approaches for its construction, such as OECD document (1998). This document generally deals with the construction of CLIs while using Phase Average Trend method to estimate long-term trend of considered economic indicators' time series and provides the calculation of the CLI for the United States.

According to the OECD methodology, CLIs are calculated for 33 OECD countries, 6 non-member countries (economies) and 8 aggregated zones on monthly basis. Table 1 shows the 5 selected countries with information on how long after the end of the reference quarter they publish the flash estimates of quarterly GDP and what is the experience with the composite indicators' construction except those calculated by OECD.

Table 1 Delays in the Transmission of the First GDP Release and Experience with Composite Indicators in Selected Countries

Country	Delay in days	Experience With Composite Indicators
Sweden	35	Only CLIs by OECD are constructed.
Austria	45	CLI constructed on monthly basis using real gross value added as a reference series. 13 indicators take part in the CLI from 91 indicators analysed.
Germany	45	Analysis of performance of leading indicator forecasts during financial crisis and performance of single and pooled leading indicators during pre-crisis and crisis period.
Italy	45	Analysis of 183 time series relevant to Italian economy on monthly basis. Combining of NBER methods and techniques of cyclical analysis.
Poland	61	Using of linear and non-linear dynamic factor modelling approaches. Predictive accuracy is confined to the in-sample-fit of the models.

Note: CIs are not designed for the purpose of the flash estimates of GDP in any of the selected countries.
Source: Eurostat; Altissimo, Marchetti, Oneto (2000); Bandholz (2005); Bierbaumer-Polly (2010); Drechsel, Scheufele (2010); own construction

2 DATA AND METHODOLOGY

The main core of this paper is to analyse relationships between appropriate and relevant time series consisting of both confidence indicators obtained from business tendency surveys, economic indicators and the cyclical component of GDP obtained from the quarterly GDP time series. The analysis is divided into two parts. The first part deals with development of time series on a visual basis, second part with co-integration analysis performed to identify type of relationship. In case of proving long-term relationship, EC model given by the formula (1) will be constructed. In the opposite case, VAR model (of size l) given by formula (2) depicting short-term relationships will be constructed. If it is proved that the considered indicator sufficiently enough explains development of GDP, the indicator would be classified as a candidate to join the composite indicator, either leading or coincidence.

$$\Delta X_t = \phi_0 + \Omega D_t + \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_{p-1} \Delta X_{t-p+1} + \Pi X_{t-p} + a_t, \tag{1}$$

$$X_t = \phi_0 + \Omega D_t + \phi_1 \Delta X_{t-1} + \ldots + \phi_p \Delta X_{t-p} + a_t, \tag{2}$$

where: $\Gamma_i = -(I_l - \phi_1 - \ldots - \phi_i)$ for $i = 1, \ldots, p-1$ and $\Pi = -(I_t - \phi_1 - \ldots - \phi_p)$ are parametric matrices containing information about relationships among processes;
ϕ_0 stands for constants, D_t stands for deterministic component and $\{a_t\}$ is the Gaussian white noise process of size l.

2.1 Selection of appropriate data and its adjustments

2.1.1 Indicators of business tendency surveys

Relatio nships of confidence indicators' time series from business tendency surveys are analysed in the form of business cycle balances defined by the Czech Statistical Office (2012) and time series of quarterly GDP (at constant 2005 prices, seasonally adjusted) cyclic component (after Hodrick-Prescott filter application) are expressed as deviations from the trend (in %). Given the data available, chosen period is from 1st quarter of 2003 to 2nd quarter of 2012. It is a period characterized by initial high economic growth that went into an economic decline due to the economic and financial crisis in 2009.

Because of the quarterly estimate of GDP at chain-linked prices of 2005 being available since 1996 and confidence indicators in manufacturing, construction and trade being available even from 1993, the series starting in 1996 (in terms of GDP) were experimentally analysed with a higher degree of assumption to prove long-term relationship, unlike the shorter ones (from 2003) but that still remain crucial to this contribution. Reliability and usefulness of the estimates of enterprises and resulting aggregated indicators are discussed by Jílek, Pecáková and Vojta (2005).

As indicators of business tendency surveys are available in monthly values, it was necessary to convert them to quarterly values by using the chronological weighted averages to permit comparison with quarterly values of the cyclical component of the GDP. The disadvantage may be a loss of information that monthly data include. Jeřábková (2010) states that other complications include the fact that the GDP by sector calculation consists of gross value added of these sectors (including net taxes on products) but the questions in business tendency surveys concern e.g. aggregate demand or economic situations and not the gross value added development, thereby commensurability of both indicators is limited.

2.1.2 Economic indicators

It is appropriate to explore other candidates for the target leading and coincidence composite indicator respectively in addition to confidence indicators for the optimization of composite indicators. There are three types of economic indicators with regard to the development of GDP and its cyclical component respectively.

The first group is represented by the leading indicators. Their task is to predict turning points in the business cycle and they are considered to be the most important group. It is clear that the choice of specific indicators is a subjective issue but it is essential to comply with certain criteria of its selection, such as simple and timely availability, high frequency of detection and using indicators that are not subject to methodology changes. In the narrowest definition among these indicators I decided to classify building permits, the number of new contracts development and stock market index. Some authors include also the Industrial Production Index in their works. For this analysis the following indicators were chosen: building permits, new contracts in the construction, new domestic contracts in manufacturing, new contracts in manufacturing from abroad, Industrial Production Index and stock market index PX (all of them in the period from 1st quarter 2003 to 2nd quarter 2012).

The second group consists of so-called coincidence economic indicators with its goal to confirm or refute the actual course of the economic cycle. Their advantage is the fact that data are available before the estimates of GDP, although both are related to the same period. Again, the choice of these indicators is subjective. Regarding the data availability and assumptions of its development the used indicators are the unemployment rate, real GDP and the index of producer prices. Since I believe that it is not correct enough to include any component directly related to GDP into the composite indicator reflecting the development of GDP, I decided to include the following indicators to the analysis: unemployment rate, index of agricultural producer prices, index of industrial producer prices, index of construction prices and index of market services prices (all of them in the period from 1st quarter 2005 to 2nd quarter 2012).

The third group includes lagged indicators used to verify the course of economic growth backwards - consumer price index, money supply and retail sales. These indicators are not included in this contribution.

2.2 Visual analysis of selected indicators

Prior to the analysis of time series in terms of existing methods, the visual analysis was called being a good starting point for getting to know time series used with respect to its development and possible connection with the investigated business cycle.

The construction of line diagram represents the key outcome and recommendation used by Czesaný, Jeřábková (2009) as well. This diagram clearly and unequivocally helps to find the location of turning points and the prevailing trend of the time series. It is also useful to combine identifiable information from a diagram with the calculated correlation coefficient between the assessed series and the number and the business cycle and to assess whether it makes sense to assign an indicator to further analysis.

2.3 Co-integration analysis as a tool of relationship analysis

Co-integration analysis has become a relatively new tool for the analysis of the time series relationship. Arlt (1997) states that time series are co-integrated if the deflection of time series' direction is only short-term, fades away over time and there is a limit that cannot be exceeded. Then it can be said that time series are located in equilibrium representing a state that the system is constantly attracted to. It is important to distinguish between stationary and non-stationary time series for analysing the time series. Co-integration is an attribute needed to perform meaningful relationships analysis among time series. For more details see Arlt and Arltová (2009).

2.4 Construction of composite indicators

Composite indicator represents an indicator composed of partial indicators of the economic cycle. This reflects the development of the economies much better than individual indicators considered separately. However, selection of the sub-indicators is not random. It is based both on its economic significance, relevance value, prediction capability and on their degree of correlation with the business cycle and even on the resulting relationship between the business cycle and these indicators for the purpose of this paper. Composition of the composite indicators in each country differs due to the significance of various indicators considered for the given economy.

Generally, there are three groups of indicators formed on the basis of its relationship to economic development. It includes leading indicators designed to predict turning points of the business cycle. Furthermore, there are coincidence indicators ordered to confirm or refuse the position of the economy and the last group are lagging indicators (this paper does not deal with them), that verify the current development of the business cycle. Tkáčová (2012) provides an overview of composite indicators' creation approaches.

3 RESULTS

This chapter introduces the most relevant results of the analysis that was at first performed for the relationships between business cycle and business tendency surveys' confidence indicators, as well as for the relationships between the business cycle and economic indicators preceding this cycle and indicators developing coincidently with the business cycle.

3.1 Relationships between business cycle and business tendency survey's indicators

The main finding is the fact that statistically significant dependence of GDP on all confidence indicators measured by business tendency survey was proved. In the analysis of its dependence on all these indicators together (except the confidence indicator in manufacturing because of its stationarity and except for the aggregate confidence indicator because of the duplicity) it is shown that GDP depends on its lagged value, on confidence indicators in trade, in services, in construction and consumer confidence indicator's lagged value. Although any long-term relationship was not shown, it can be stated that aggregate confidence indicator is an appropriate sub-indicator for CLI. VAR models for all partial indicators showed that statistically significant dependence exists between GDP and the corresponding number of partial confidence indicators, as well as in the case of examining the relationship between GDP and all these sub-indicators together, where the relationships were identified too, although only short-term. It is definitely caused by the relatively short time series and it can be assumed that there will be the evidence of long-term relationships in the future.

3.2 Relationships between business cycle and leading and coincidence economic indicators

From selected indicators which precede business cycle only one will not be included in the CLI, namely New Contracts in the Construction, as between its time series and GDP series have not been identified even any statistically significant short-term relationships. In the analysis of the relationship between GDP and all leading indicators its series were nonstationary, this indicator explains the development of GDP (albeit temporarily) with high, 5-quarter lag. Another such indicator is the Building Permit that relatively poorly explains GDP development (or its first difference). There is also very low correlation coefficient indicating very weak indirect linear dependence between the range of GDP values and range of Building Permits indicator's values.

Using the VAR model, short-term relationships between GDP and sectional coincidence economic indicators were modelled and although Market Services and Construction Work Price Indices did not seem to be appropriate for participation in the CCI by visual analysis, co-integration analysis refuted its ability to explain GDP development and therefore they will be included in the composite indicator. Composite indicators were constructed by 3 basic steps – normalization, weighting and aggregation. The resulting CLI (see Figure 1) was constructed with equal weights due to its better relationship to the business cycle, while CCI (see Figure 2) was constructed with different weights (derived by the correlation coefficient value between business cycle series and the relevant economic indicator's series) due to the same reason. Overview of all selected indicators for composite indicators' construction including used weights is represented by Table 2 in the case of CLI and by Table 3 in the case of CCI.

Table 2 Overview and Weights of Selected Economic Indicators Entering CLI

Leading Composite Indicator	Weight
Aggregate Confidence Indicator	0.2
New Contracts from Domestic Manufacturing	0.2
New Contracts from Abroad Manufacturing	0.2
Industrial Production Index	0.2
Stock Market Index PX	0.2

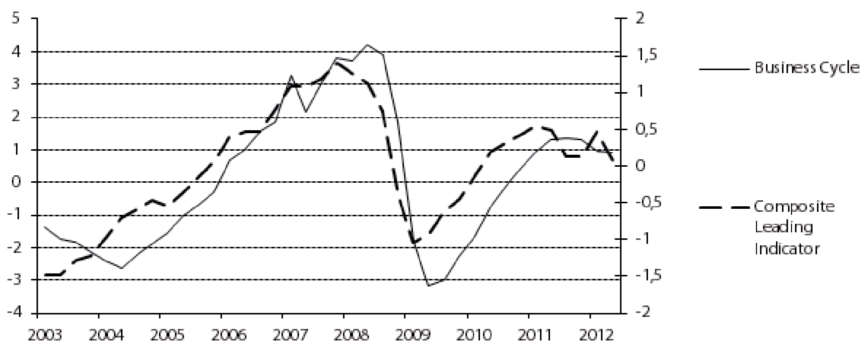
Source: Own calculation

Table 3 Overview and Weights of Selected Economic Indicators Entering CCI

Coincidence Composite Indicator	Weight
Unemployment Rate	0.3287
Agricultural Producer Price Index	0.3434
Manufacturing Producer Price Index	0.3099
Construction Work Price Index	0.0015
Market Services Price Index	0.0165

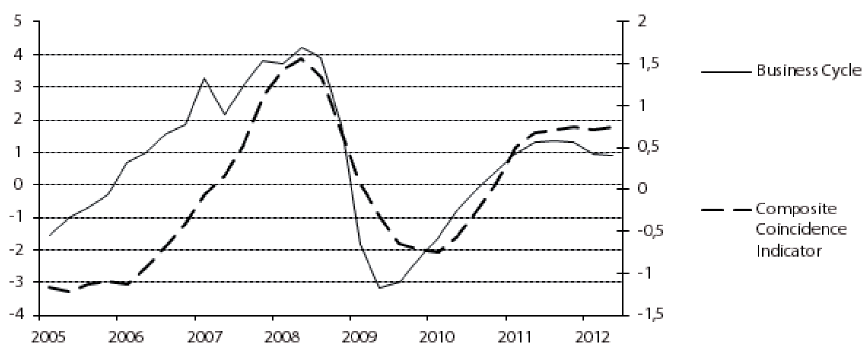
Source: Own calculation

Figure 1 Development of the Composite Leading Indicator (equal weights) and Business Cycle (in % of trend)



Source: Own construction

Figure 2 Development of the Composite Coincidence Indicator (different weights) and Business Cycle (in % of trend)



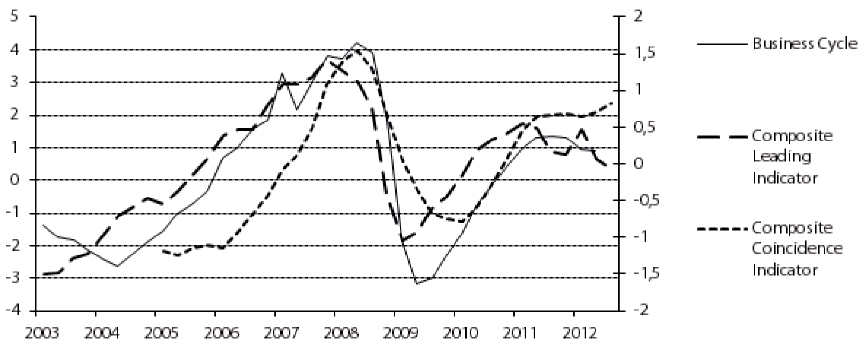
Source: Own construction

4 VERIFICATION OF THE RESULTING COMPOSITE INDICATORS REGARDING THE ACTUAL DEVELOPMENT OF THE BUSINESS CYCLE

All previous calculations and analyses included periods with beginnings chosen regarding the data availability. The last period was the second quarter of 2012. During writing this paper, monthly and quarterly values of considered coincidence and leading indicators of third quarter of 2012 were published and it allows usage of developed composite indicators to verify the quality of the estimation of quarterly GDP change for the third quarter of 2012.

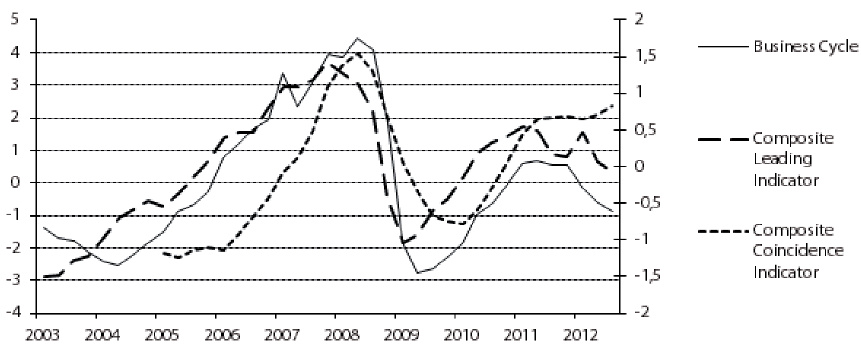
The value of the CCI is 0.833 in the third quarter of 2012 and 0.702 in the second quarter of 2012. From these values it is possible to conclude that GDP should increase quarter to quarter. If we look at the CLI and assuming that its value outpaces GDP usually about 2 quarters, this GDP growth assumption is confirmed. CLI's value is 0.063 in the second quarter of 2012 and -0.064 in the third quarter of 2012, so it is apparent that in the fourth quarter of 2012 and in the first quarter of 2013 GDP should decline. For illustration, see Figure 3 including business cycle, CLI and CCI.

Figure 3 Development of the Composite Leading Indicator, Composite Coincidence Indicator and Business Cycle (in % of trend)



Source: Own construction

Figure 4 Development of the Composite Leading Indicator, Composite Coincidence Indicator (right axis) and Business Cycle (in % of trend, left axis) including 3rd quarter of 2012



Source: Own construction

On 15th November 2012 (45 days after the end of the third quarter of 2012) a preliminary estimate of quarterly GDP, that declined by 0.3% quarter to quarter, was published by the Czech Statistical Office. Its seasonally adjusted value is 893.973 million CZK and value of the cyclical component expressed as a deviation from the trend is 0.699%. Figure 4 captures the evolution of business cycle including third quarter of 2012 and the development of both composite indicators.

According to these results, the assumption of moderate business cycle growth is refuted. In the first quarter, CLI showed a blip that indicated business cycle could increase. It was also supported by the coincidence composite indicator's value increase but the reality consisted in the decrease of the business cycle. In conclusion, the constructed composite indicators are needed to be approached with caution. It is required to follow the individual economic indicators' (in the composite indicators entered) development and subject these composite indicators to regular revisions.

CONCLUSION

The issue of quarterly estimate of GDP is a relatively wide range of possible approaches to achieve this goal. Perhaps the biggest challenge is the lack of long time series that would certainly prove the presence of long-term relationships between GDP and economic indicators analysed. Another issue is the choice of economic indicators that vary in authors different approaches. For example, some indicators included are directly related or taking part in actual GDP, while this paper deals only with the basic economic indicators that can be found in macroeconomic textbooks and regarding data availability and timeliness. In this paper, majority of selected indicators affects manufacturing hence manufacturing has relatively important position in the Czech economy since there is more than one third of gross value added created in this industry.

In relation to the form of this analysis, it is necessary to emphasize the need of regular revisions of these composite indicators and the need of updating the weights used. However, it is necessary to treat these indicators with sufficient margin and to monitor the development of sub-indicators as a complementary source of data.

This contribution should serve rather as starting a new approach to the estimation of the development of quarterly GDP (using time series methods) that has to be further expanded and improved in the issues mentioned above. For further research it is also offered, in addition to the identifying the direction of GDP development, its quantification with subsequent validation and comparison with real development as well.

References

ALTISSIMO, F., MARCHETTI, D. J., ONETO, G. P. The Italian Business Cycle: Coincident and Leading Indicators and Some Stylized Facts. *Giornale degli Economisti e Annali di Economia*, 2000, Vol. 60, No. 2, pp. 147–220. ISSN 0017-0097.

ARLT, J., ARLTOVÁ, M. *Ekonomické časové řady* (Economic time series). Prague: Professional Publishing, 2009. ISBN 978-80-86946-85-6.

ARLT, j. Kointegrace v jednorovnicových modelech (Cointegration in single-equation models). *Politická ekonomie,* 1997, Vol. 45, No. 5, pp. 733–746. ISSN 0032-3233.

BANDHOLZ, H. New Composite Leading Indicators for Hungary and Poland. *IFO Working Papers,* 2005, No. 3, 32 p.

BIERBAUMER-POLLY, J. Composite Leading Indicator for the Austrian Economy: Methodology and „Real-Time" Performance. *WIFO Working Papers,* 2010, No. 369, 44 p.

CZECH STATISTICAL OFFICE. *Konjunkturální průzkum* (Business cycle survey). [online]. [cit. 27.10.2012]. <http://www.czso.cz/csu/redakce.nsf/i/konjunkturalni_pruzkum>.

CZESANÝ, S., JEŘÁBKOVÁ, Z. Metoda konstrukce kompozitních indikátorů hospodářského cyklu pro českou ekonomiku (Method of construction of composite indicators of the business cycle for the Czech economy). *Statistika*, 2009, Vol. 46, No. 1, pp. 21–31. ISSN 1804-8765.

DRECHSEL, K., SCHEUFELE, R. Should We Trust in Leading Indicators? Evidence From Recent Recession. *IWH Discussion Papers*, Halle Institute for Economic Research, 2010, No. 10, 48 p. ISSN 1860-5303.

EUROSTAT. Flash estimation of the quarterly Gross Domestic Product for the euro-zone and the European Union. E*urostat methodology.* Luxembourg: Office for Official Publications of the European Communities, 2003. ISSN 1725-0048.

FISCHER, J. *Konstrukce čtvrtletních předběžných odhadů vývoje hrubého domácího produktu s ohledem na charakter a vypovídací schopnosti oficiálních odhadů* (Construction of preliminary quarterly estimates of the GDP considering the character and explanatory power of the official estimates). Disertační práce v oboru "Statistika" (Dissertation in "Statistics"). Prague: University of Economics (FIS, Department of Economic Statistics), 2005, 107 p.

FISCHER, J. a kol: *Řady konjunkturních sald jako vysvětlující proměnné řad relativních změn HDP ČR* (The series of the business cycle balances as explanatory variables of the relative changes in the GDP of CR). Průběžná zpráva o plnění dílčího úkolu v roce 2002 (Continuous report about the performance of a subtask in 2002), Prague: University of Economics (FIS), 2002.

JEŘÁBKOVÁ, Z. *Konjunkturální indikátory a jejich vazba na hospodářský vývoj v době zpomalení české ekonomiky* (Business cycle indicators and their link to the economic development during deceleration of the Czech economy). [online]. Prague: Czech Statistical Office, 2010. [cit. 22.11.2012]. <http://www.czso.cz/csu/csu.nsf/informace/ckta19072010.doc>.

JÍLEK, J., PECÁKOVÁ, I., VOJTA, M. Konjunkturální průzkumy ČSÚ a predikce (Business cycle surveys of the CZSO and predictions). *Statistika*, 2005, Vol. 42, No. 5, pp. 395–425. ISSN 0322-788X.

OECD. *OECD Composite Leading Indicators – a tool for short-term analysis.* OECD, Statistics Directorate, 1998 [online]. [cit. 25.1.2014]. <http://www.oecd.org/std/leading-indicators/15994428.pdf>.

POŠTA, V., VALENTA, V. Composite Leading Indicators Based on Business Surveys: Case of the Czech Economy. *Statistika,* 2011, Vol. 48, No. 1, pp. 12–18. ISSN 0322-788X.

TKÁČOVÁ, A. Kompozitný predstihový indikátor hospodárskeho cyklu českej ekonomiky (The composite advancer business cycle indicator of the Czech economy). *Politická ekonomie*, 2012, No. 5, pp. 590–611. ISSN 0032-3233.

Application of Stochastic Index Numbers in Inflation Measurement – the Case of Poland

Jacek Białek[1] | *University of Lodz, Lodz, Poland*

Abstract

The stochastic approach is a specific way of viewing index numbers, in which uncertainty and statistical properties play a central role. This approach, applied to the prices, treats the underlying rate of inflation as an unknown parameter that has to be estimated from the individual prices. Thus, the stochastic approach provides the whole probability distribution of inflation. In this paper we present and discuss several basic stochastic index numbers. We propose a simple stochastic model, which leads to a price index formula being a mixture of the previously presented specifications. We verify the considered indices on a real data set.

Keywords	***JEL code***
Price indices, stochastic index numbers, price index theory	*E17, E21, E30*

INTRODUCTION

The weighted price index is a function of a set of prices and quantites of the considered group of N commodities comming from the given moment t and the basic moment s. In reality, the price index formula is a quotient of some random variables and thus, it is really difficult to construct a confidence interval for that formula. The so called new stochastic approach (NSA) in the price index theory gives a solution for the above-mentioned problem. Within this approach, a price index is a regression coefficient (unknown parameter[2] θ) in a model explaining price variation. Having estimated sampling variance of the estimator ($\hat{\sigma}_{\theta}^{2}$) we can build the $(1 - \alpha)$ confidence interval[3] as $\hat{\theta} \pm t_{1-\alpha/2,n-1}\hat{\sigma}_{\theta}$, where n is the sample size and $t_{1-\alpha/2,n-1}$ is the $100(1-\alpha/2)$ percentile of the t distribution with $n-1$ degrees of freedom (see von der Lippe (2007)). The individual prices are observed with error and the problem is a signal-extraction one of how to combine noisy prices so as to minimize the effects of measurement errors. Under certain assumptions, the stochastic approach leads to known price index formulas (such as Divisa, Laspees, etc.), but their foundations differ from the classical deterministic approach. Within this approach we can also

[1] University of Lodz, Chair of Statistical Methods, 90-255 Lodz, 3/5 POW Street, Poland. E-mail: jbialek@uni.lodz.pl.

[2] In the next part of the paper we consider only the least squares and maximum likelihood estimator for θ.

[3] We can build this confidence interval under the additional assumption that the residuals are normally distributed.

obtain some new price index formulas having some desired economical and statistical properties (Clements et al. (2006)).

The stochastic approach originated in the work of Jevons (1863, 1869) and Edgeworth (1887, 1888, 1889). Aldrich (1992) attributes the introduction of the term "stochastic" in this context to Frisch (1936), and it was adopted by Allen (1975), to describe Edeworth's analysis. The recent resurrection of the stochastic approach to index number theory is due to Balk (1980), Clements and Izan (1981, 1987), Bryan and Cecchetti (1993) and Selvanathan and Prasada Rao (1992). This literature is still expanding and has been the subject of a book by Selvanathan and Prasada Rao (1994), who emphasise the versatility and the usefulness of the stochastic approach. Although some papers have critical tone (see for example Diewert (1995)), some other and more recent papers extend this approach in new directions (see Diewert (2004, 2005), Prasada Rao (2004)). In this paper we present and discuss only some basic stochastic index numbers. We propose a simple stochastic model, which leads to a price index formula being a mixture of the previously presented specifications.

1 STOCHASTIC INDEX NUMBERS IN INFLATION MEASUREMENT

The main attraction of the stochastic approach over competing approaches to the index number theory is its ability to provide confidence intervals for the estimated inflation rates:

"Accordingly, we obtain a point estimate of not only the rate of inflation, but also its sampling variance. The source of the sampling error is the dispersion of relative prices from their trend rates of change -- the sampling variance will be larger when the deviations of the relative prices from their trend rates of change are larger. This attractive result provides a formal link between the measurement of inflation and changes in relative prices." (Clements and Izan (1987), p. 339)

There are many directions and stochastic models in the field of the stochastic approach. To make the exposition of stochastic index numbers as clear as possible, we concentrate on the simplest possible cases. Let $Dp_{i,t} = \ln p_{i,t} - \ln p_{i,t-1}$ be the log-change in price of commodity i ($i = 1,2,\ldots, N$) from year $t-1$ to t. Suppose that each price change is made up of a systematic part that is common to all prices (θ_t) and a random component $\varepsilon_{i,t}$,

$$Dp_{i,t} = \theta_t + \varepsilon_{i,t}, \tag{1}$$

where we assume that $E(\varepsilon_{i,t}) = 0$ and thus $E(Dp_{i,t}) = \theta_t$. We can see that the parameter θ_t is interpreted here as the common trend in all prices, or the underlying rate of inflation. Let all $\varepsilon_{i,t}$ have variances and covariances of the form $\hat{\sigma}^2_{ij,t}$ and let $\Sigma_t = [\hat{\sigma}^2_{ij,t}]$ be the corresponding $N \times N$ covariance matrix. Under above significations we can write (1) in vector form as:

$$Dp_t = \theta_t u + \varepsilon_t, \tag{2}$$

where $Dp_t = [Dp_{i,t}]'$, $u = [1,\ldots,1]'$, $\varepsilon_t = [\varepsilon_{i,t}]'$ are all $N \times 1$ vectors.

Using the generalized least squares method for estimating θ_t we obtain the BLUE estimator as follows (see Clements et al. (2006)):

$$\hat{\theta}_t = (u'\Sigma_t^{-1}u)^{-1}u'\Sigma_t^{-1}Dp_t, \tag{3}$$

with variation:

$$\hat{\sigma}^2_{\theta_t} = (u'\Sigma_t^{-1}u)^{-1}. \tag{4}$$

The presented formulas (3) and (4) have a general form and in the remaining part of the paper we consider some special cases of this model. Let us notice that $\varepsilon_{i,t}$ is interpreted as the change in the i – th relative price. Let us suppose that $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$ are independent (for $i \neq j$) and

$$\sigma^2_{ii,t} = \frac{\lambda^2_t}{w_{i,t}}, \tag{5}$$

where λ_t is a parameter independent of i and $w_{i,t}$ is the i – th budget share, with $q_{i,t}$ the quantity consumed of i -th good in the year t, namely:

$$w_{i,t} = \frac{p_{i,t} q_{i,t}}{\sum_{i=1}^{N} p_{i,t} q_{i,t}}. \tag{6}$$

The assumption (5) means that the variance of $\varepsilon_{i,t}$ is inversely proportional to the corresponding budget share $w_{i,t}$. There are several justifications for the specification (5). One of them (see Clements et al. (2006)) can be written as follows: since a commodity absorbs a large part of the of the overall economy (its budget share rises), there is less scope for its relative price to vary as there is simply a lesser amount of other goods against which its price can change. In other words the variance of a large good is smaller than the variances of other goods. It can be easily shown that then we get (see Clements et al. (2006)):

$$\Sigma_t = \lambda^2_t W^{-1}_t, \tag{7}$$

where $W_t = diag[w_{1,t}, w_{2,t}, \ldots, w_{N,t}]$.

From (3), (4) and (7) we obtain[4] (see also von der Lippe (2007)):

$$\hat{\theta}^I_t = \sum_{i=1}^{N} w_{i,t} Dp_{i,t}, \tag{8}$$

$$\hat{\sigma}^2_{\hat{\theta}^I_t} = \frac{1}{N-1} \sum_{i=1}^{N} w_{i,t} (Dp_{i,t} - \hat{\theta}^I_t)^2. \tag{9}$$

In other words, the estimator $\hat{\theta}^I_t$ of the underlying rate of inflation is a budget-share weighted average of the N price log-changes. It makes intuitive sense. Moreover, we can notice that $\exp(\hat{\theta}^I_t)$ is a logarithmic Paasche price index, and if we use as weights the arithmetic average of the observed budget shares in years $t-1$ and t, we obtain in (8) the Divisia price index, also known as the Törnqvist (1936)-Theil (1967) index, that has many of desirable properties.

As it was already mentioned, Diewert (1995) criticizes the stochastic approach. One of his remarks is that the variance assumptions are not consistent with the facts. Diewert argues that equation (5) is not in line with observed behavior of prices.[5] Some authors reject this specification (see Clements and Izan (1987)) but let us notice, that variance specification (5) is just one of multitude of possibilities. In the paper by Clements et al. (2006) authors give three other specifications to show how the stochastic approach deals with different specifications of Σ_t – case I: prices are independent (Σ_t is a diagonal matrix with elements $\sigma^2_{11,t}, \sigma^2_{22,t}, \ldots, \sigma^2_{NN,t}$); case II: prices have a common variance σ^2_t and a common correlation coefficient ρ_t at time t ($\Sigma_t = \sigma^2_t[(1-\rho_t)I + \rho_t uu']$, where I is an identity matrix); case III: $\Sigma_t = D_t(I + \lambda_t) D_t$, where D_t is a diagonal matrix with the standard deviation of N prices on the main diagonal, $\lambda_t = [\lambda_{ij,t}]$ is an $N \times N$ symmetric matrix with diagonal elements zero and (i,j) – th off-diagonal element the relevant correlation, it means $\lambda_{ij,t} = \sigma^2_{ij,t} / (\sigma_{ij,t}\sigma_{jj,t})$.

The afore-mentioned authors show that depending on the case we get:

$$\text{case I: } \hat{\theta}^I_t = \sum_{i=1}^{N} w^I_{i,t} Dp_{i,t}, \quad \hat{\sigma}^2_{\hat{\theta}^I_t} = \frac{1}{\sum_{i=1}^{N} \sigma^{-2}_{ii,t}}, \text{ where } w^I_{i,t} = \frac{\sigma^{-2}_{ii,t}}{\sum_{i=1}^{N} \sigma^{-2}_{ii,t}};$$

[4] To distinguish estimators coming from different models we use the following notation: $\hat{\theta}^I_t$, $\hat{\theta}^{II}_t$,…

[5] Diewert (1995) gives the following example: food has a big share while energy has a small share and the volatility of price components is simply not highly correlated with the corresponding expenditure shares.

case II: $\hat{\theta}_t^{II} = \frac{1}{N}\sum_{i=1}^{N} Dp_{i,t}$, $\hat{\sigma}^2_{\hat{\theta}_t^{II}} = \sigma_t^2[\rho_t + \frac{1-\rho_t}{N}]$;

case III: $\hat{\theta}_t^{III} = \sum_{i=1}^{N} w_{i,t}^{III} Dp_{i,t}$, $\hat{\sigma}^2_{\hat{\theta}_t^{III}} = \frac{1}{\sum_{i=1}^{N}(\sigma_{ii,t}^{-2} - \lambda_{.i}^{*})}$,

where $w_{i,t}^{III} = \frac{\sigma_{ii,t}^{-2} - \lambda_{.i,t}^{*}}{\sum_{i=1}^{N}(\sigma_{ii,t}^{-2} - \lambda_{.i,t}^{*})}$

and $\lambda_{i,t}^{*}$ is the sum of elements in the i – th row of the matrix $\lambda_t^{*} = D_t^{-1}\lambda_t D_t^{-1}$,

namely $\lambda_{.i,t}^{*} = \sum_{j=1}^{N} \lambda_{ij,t}^{*}$.

As we can see in the first case the estimated rate of inflation is still a weighted average of the price changes, but now the weights are proportional to the reciprocals of the variances of the respective relative prices. Obviously, the weights are positive and have a unit sum. In the second case, the estimated rate of inflation is an unweighted average of price changes, while its variance is increasing in the common correlation ρ_t. In this case, if prices are independent we obtain $\hat{\sigma}^2_{\hat{\theta}_t^{II}} = \sigma_t^2/N$ and if prices are perfectly and positively correlated we have $\hat{\sigma}^2_{\hat{\theta}_t^{II}} = \sigma_t^2$. In case III, which is the most realistic, the estimated rate of inflation is again a weighted average of price changes[6] but now the weights $w_{i,t}^{III}$ are related to the variances and covariances of the relative prices. The fraction $w_{i,t}^{III}$ is larger when the i – th variance is lower and the i – th relative price is less correlated with the others. In cases II and III the value of the estimator does not depend on the budget share. Other specifications of the covariance matrix are clearly possible (see Crompton (2000)) and we propose one of them in the next part of the paper. Although the form of the matrix Σ_t determines the final results, still the main idea is to think of the rate of inflation as the underlying common trend in prices. As we can notice, in the presented stochastic models this trend is estimated by a type of a mean of the considered N price changes.

2 A BASIC MODEL AND A NEW PRICE INDEX FORMULA

Let us assume the following specification[7] of the matrix Σ_t:

$$\Sigma_t = D_t(I - \lambda_t)^{-1} D_t W_t^{-1}, \tag{10}$$

where D_t is diagonal matrix with the standard deviations of N relative prices on the main diagonal, $\lambda_t = [\lambda_{ij,t}]$ is an $N \times N$ symmetric matrix with diagonal elements zero and (i,j) – th off-diagonal element the relevant correlation[8] $\lambda_{ij,t} = \sigma_{ij,t}^2/(\sigma_{ii,t}\,\sigma_{jj,t})$ and $W_t = diag[w_{1,t}, w_{2,t},..., w_{N,t}]$ is an $N \times N$ diagonal matrix. The following theorem is true.

Theorem 1

In the stochastic model described by (1) with the corresponding covariance matrix defined by (10) we obtain the following estimator of the rate of inflation[9] and its variation:

---

6 To be precise the formula describing the estimator in case III is an approximation, since it holds that $(I - \lambda_t)^{-1} \approx I + \lambda_t$.

7 The specification (10) is similar to the specification presented previously as case III, namely $\Sigma_t = D_t(I + \lambda_t)D_t$. In fact, from the known result that $(I - \lambda_t)^{-1} = I + \lambda_t + \lambda_t^2 + ...$ for small elements of λ_t we have $(I - \lambda_t)^{-1} \approx I + \lambda_t$. The last component, the matrix W_t, corresponds to the budget share model (BSM – see von der Lippe (2007)), also presented previously. In other words, the present model is some kind of mixture of the earlier models.

8 We assume here the realistic scenario that prices are correlated. Otherwise, we should take $\Sigma_t = D_t^2 W_t^{-1}$.

9 We still use the generalized least squares method for estimating.

$$\hat{\theta}_t^* = \sum_{i=1}^{N} w_{i,t}^* Dp_{i,t}, \tag{11}$$

$$\hat{\sigma}_{\hat{\theta}_t^*}^2 = \frac{1}{\sum_{i=1}^{N} w_{i,t}(\sigma_{ii,t}^{-2} - \lambda_{.i,t}^*)}, \tag{12}$$

where:

$$w_{i,t}^* = \frac{w_{i,t}(\sigma_{ii,t}^{-2} - \lambda_{.i,t}^*)}{\sum_{i=1}^{N} w_{i,t}(\sigma_{ii,t}^{-2} - \lambda_{.i,t}^*)}, \tag{13}$$

and $\lambda_{i,t}^*$ described as in case III (see section 1).

Proof
Firstly, from (10) we obtain:

$$\Sigma_t^{-1} = W_t D_t^{-1}(I - \lambda_t) D_t^{-1}, \tag{14}$$

and thus, we have:

$$[u'\Sigma_t^{-1}u]^{-1} = [u'W_t D_t^{-1}(I - \lambda_t) D_t^{-1}u]^{-1} = [u'W_t(D_t^{-2}u - D_t^{-1}\lambda_t D_t^{-1}u)]^{-1} =$$

$$= [\sum_{i=1}^{N} w_{i,t}(\sigma_{ii,t}^{-2} - \sum_{j=1}^{N} \sigma_{ii,t}^{-1}\lambda_{ij,t}\sigma_{jj,t}^{-1})]^{-1} = [\sum_{i=1}^{N} w_{i,t}(\sigma_{ii,t}^{-2} - \lambda_{.i,t}^*)]^{-1}. \tag{15}$$

The second part of the right-hand side of the equation (3) is as follows:

$$u'\Sigma_t^{-1}Dp_t = u'[W_t D_t^{-1}(I - \lambda_t) D_t^{-1}]Dp_t = u'[W_t(D_t^{-2} - D_t^{-1}\lambda_t D_t^{-1})]Dp_t =$$

$$= u'[W_t(D_t^{-2} - \lambda_t^*)]Dp_t = \sum_{i=1}^{N} w_{i,t}(\sigma_{ii,t}^{-2} - \lambda_{.i,t}^*)Dp_{i,t}. \tag{16}$$

From (3), (15) and (16) we obtain:

$$\hat{\theta}_t^* = (u'\Sigma_t^{-1}u)^{-1}u'\Sigma_t^{-1}Dp_t = \sum_{i=1}^{N} \frac{w_{i,t}(\sigma_{ii,t}^{-2} - \lambda_{.i,t}^*)}{\sum_{i=1}^{N} w_{i,t}(\sigma_{ii,t}^{-2} - \lambda_{.i,t}^*)} Dp_{i,t} = \sum_{i=1}^{N} w_{i,t}^* Dp_{i,t}. \tag{17}$$

Let us notice that from (4) and (15) we get directly the variation of the estimator :

$$\hat{\sigma}_{\theta_t^*}^2 = (u'\Sigma_t^{-1}u)^{-1} = \frac{1}{\sum_{i=1}^{N} w_{i,t}(\sigma_{ii,t}^{-2} - \lambda_{.i,t}^*)}. \tag{18}$$

Remark
As we can see the estimated rate of inflation (11) with weights described by (13) is still a weighted arithmetic mean of the price log-changes, where the weights are proportional to the reciprocals of the variances of the relative prices, proportional to the budget-shares and it also takes into account correlations among prices. In the next part of the paper (see the empirical study) we compare results obtained by using estimators $\hat{\theta}_t^{III}$ and $\hat{\theta}_t^*$.

3 EMPIRICAL STUDY

In our empirical illustration of the presented measures of inflation we use monthly data[10] on price indices of consumer goods and services in Poland for the time period I 2010–XII 2012 (36 observations). The weights $w_{i,t}$ also are taken from data published by the Central Statistical Office.[11] The calculated standard deviations of considered relative prices and their correlations for each considered year are presented in (respectively) Table 1 and Table 2. The estimated rates of inflation for years: 2010–2012 with the corresponding variations and confidence intervals are presented in Table 3.

Table 1 Standard deviations of the log-change prices of the considered goods and services in Poland

Year	Standard deviations											
	X1	*X2*	*X3*	*X4*	*X5*	*X6*	*X7*	*X8*	*X9*	*X10*	*X11*	*X12*
2010	0.0131	0.0173	0.0079	0.0038	0.0043	0.0007	0.0240	0.0037	0.0091	0.0024	0.0041	0.0042
2011	0.0140	0.0051	0.0154	0.0044	0.0053	0.0123	0.0123	0.0166	0.0042	0.0105	0.0031	0.0036
2012	0.0080	0.0050	0.0122	0.0062	0.0022	0.0125	0.0243	0.0138	0.0038	0.0106	0.0024	0.0045

Source: Own calculations

Table 2 Correlations of the considered log-change prices for years 2010–2012 in Poland

Year: 2010	*X1*	*X2*	*X3*	*X4*	*X5*	*X6*	*X7*	*X8*	*X9*	*X10*	*X11*	*X12*
X1	1.000											
X2	–0.687	1.000										
X3	0.591	–0.863	1.000									
X4	0.753	-0.451	0.476	1.000								
X5	–0.511	0.870	–0.959	–0.418	1.000							
X6	–0.132	0.244	–0.483	0.037	0.357	1.000						
X7	–0.399	0.854	–0.680	–0.052	0.770	0.233	1.000					
X8	0.324	0.030	0.147	0.655	–0.028	–0.306	0.371	1.000				
X9	–0.058	0.593	–0.547	0.246	0.676	0.086	0.830	0.692	1.000			
X10	–0.070	–0.256	–0.082	–0.166	–0.009	0.227	–0.468	–0.258	–0.202	1.000		
X11	–0.409	0.888	–0.898	–0.268	0.954	0.259	0.860	0.114	0.754	–0.233	1.000	
X12	–0.604	0.920	–0.895	–0.363	0.951	0.324	0.845	0.039	0.679	–0.115	0.924	1.000

Year: 2011	*X1*	*X2*	*X3*	*X4*	*X5*	*X6*	*X7*	*X8*	*X9*	*X10*	*X11*	*X12*
X1	1.000											
X2	–0.805	1.000										
X3	0.602	–0.518	1.000									
X4	–0.634	0.324	–0.003	1.000								
X5	–0.405	–0.002	0.256	0.828	1.000							
X6	–0.240	–0.163	0.046	0.615	0.745	1.000						
X7	–0.407	0.005	–0.194	0.717	0.610	0.630	1.000					
X8	–0.494	0.084	–0.468	0.581	0.467	0.390	0.777	1.000				
X9	0.138	-0.563	0.288	0.402	0.688	0.596	0.566	0.433	1.000			
X10	–0.608	0.176	–0.265	0.795	0.767	0.603	0.761	0.713	0.555	1.000		
X11	0.129	–0.346	0.760	0.456	0.734	0.359	0.108	0.006	0.602	0.274	1.000	
X12	–0.171	–0.245	0.472	0.690	0.930	0.582	0.482	0.341	0.766	0.635	0.882	1.000

[10] We use highly-aggregated data taking into account price indices of the following group of consumer goods and services in Poland: food and non-alkoholic beverages (*X1*), alcoholic beverages, tobacco (*X2*), clothing and footwear (*X3*), housing, water, electricity, gas and other fuels (*X4*), furnishings, household equipment and routine maintenance of the house (*X5*), health (*X6*), transport (*X7*), communications (*X8*), recreation and culture (*X9*), education (*X10*), restaurants and hotels (*X11*) and miscellaneous goods and services (*X12*).

[11] Główny Urząd Statystyczny (GUS) in Poland.

Table 2 Correlations of the considered log-change prices for years 2010–2012 in Poland continuation

Year: 2012	*X1*	*X2*	*X3*	*X4*	*X5*	*X6*	*X7*	*X8*	*X9*	*X10*	*X11*	*X12*
X1	1.000											
X2	–0.603	1.000										
X3	–0.260	0.396	1.000									
X4	–0.222	0.736	0.353	1.000								
X5	–0.155	0.563	0.580	0.871	1.000							
X6	0.129	0.510	0.344	0.770	0.827	1.000						
X7	0.051	0.596	0.514	0.791	0.846	0.943	1.000					
X8	0.056	0.160	–0.372	0.517	0.236	0.166	0.078	1.000				
X9	–0.003	0.607	0.269	0.431	0.322	0.575	0.515	–0.017	1.000			
X10	–0.143	0.603	0.441	0.829	0.737	0.715	0.709	0.315	0.501	1.000		
X11	0.101	0.022	0.734	0.245	0.538	0.460	0.568	–0.397	0.074	0.500	1.000	
X12	–0.133	0.498	0.771	0.667	0.782	0.713	0.845	–0.106	0.283	0.770	0.832	1.000

Source: Own calculations

Table 3 Values of the considered estimators of a rate of inflation, their variances and the corresponding 95% confidence intervals for years 2010–2012 in Poland

Measure	Year: 2010 (published[12] rate of inflation –0.031)		
$\hat{\theta}_t^{III}$	0.0334	$\hat{\theta}_t^{*}$	0.0325
$\hat{\sigma}^2_{\hat{\theta}_t^{III}}$	0.0129	$\hat{\sigma}^2_{\hat{\theta}_t^{*}}$	0.0023
95% confidence interval	(0.0049; 0.0620)	95% confidence interval	(0.0274; 0.0376)
Measure	**Year: 2011 (published rate of inflation –0.046)**		
$\hat{\theta}_t^{III}$	0.0405	$\hat{\theta}_t^{*}$	0.0474
$\hat{\sigma}^2_{\hat{\theta}_t^{III}}$	0.0083	$\hat{\sigma}^2_{\hat{\theta}_t^{*}}$	0.0011
95% confidence interval	(0.0220; 0.0588)	95% confidence interval	(0.0450; 0.0498)
Measure	**Year: 2012 (published rate of inflation –0.024)**		
$\hat{\theta}_t^{III}$	0.0183	$\hat{\theta}_t^{*}$	0.0239
$\hat{\sigma}^2_{\hat{\theta}_t^{III}}$	0.0061	$\hat{\sigma}^2_{\hat{\theta}_t^{*}}$	0.0009
95% confidence interval	(0.0049; 0.0317)	95% confidence interval	(0.0219; 0.0259)

Source: Own calculations

CONCLUSIONS

It is not unexpected that values of estimators $\hat{\theta}_t^{III}$ and $\hat{\theta}_t^{*}$ differ from each other and values of $\hat{\theta}_t^{*}$ are closer to the published rates of inflation, because only $\hat{\theta}_t^{*}$ and CPI[13] take into account budget shares. However, $\hat{\theta}_t^{III}$ and $\hat{\theta}_t^{*}$ have the same merit – they also take into account variances and correlations of the relative prices. Moreover, the general conclusion is that the variance of the $\hat{\theta}_t^{*}$ estimator (for each year of the research) is smaller then the variance of $\hat{\theta}_t^{III}$ and thus, the confidence intervals for $\hat{\theta}_t^{*}$ are more narrow then confidence

[12] This is an official yearly rate of inflation in Poland published by the Central Statistical Office in December of a given year. To be more precise it is a value of the general price index of consumer goods and services (December of the previous year is a base period) minus one. This value should be approximated by $\exp(\hat{\theta}_t) - 1$, but we use $\hat{\theta}_t$ as an approximation since $\exp(\hat{\theta}_t) - 1 \approx \hat{\theta}_t$ for small values of $\hat{\theta}_t$.

[13] CPI (Consumer Price Index) in Poland takes the Laspeyres form.

intervals calculated for $\hat{\theta}_t^{III}$. In particular, the published rate of inflation in Poland seems to be too small in 2010 (it equals 3,1%, when $\hat{\theta}_t^{III}$ = 3,34% and $\hat{\theta}_t^{*}$ = 3,25%) and overestimated in 2012 (it equals 2,4%, when $\hat{\theta}_t^{III}$ = 1,83% and $\hat{\theta}_t^{*}$ = 2,39%). Let us also notice that all confidence intervals for estimated rate of inflation include the value of this rate published by the Central Statistical Office in the corresponding year.

References

ALDRICH, J. Probability and Depreciation – A History of Stochastic Approach to Index Numbers. *History of Political Economy*, 1992, 24, pp. 657–687.

ALLEN, R. G. D. *Index Numbers in Theory and Practice*. New York: Macmillan Press, 1975.

BALK, B. M. A Method for Constructing Price Indices for Seasonal Commodities. *Journal of the Royal Statistical Society,* A, 1980, 143, pp. 68–75.

BRYAN, M. F., CECCHETTI, S. G. The Consumer Price Index as a Measure of Inflation. *Economic Review*, Federal Reserve Bank of Cleveland, 1993, 29, pp. 15–24.

CLEMENTS, K. W., IZAN, H. Y. A Note on Estimating Divisia Index Numbers. *International Economics Review,* 1981, 22, pp. 745–747.

CLEMENTS, K. W., IZAN, H. Y. The Measurement of Inflation: A Stochastic Approach. *Journal of Business and Economic Statistics*, 1987, 5, pp. 339–350.

CLEMENTS, K. W., IZAN, H. Y., SELVANATHAN, E. A. Stochastic Index Numbers: A Review. *International Statistical Review,* 2006, 74, 2, pp. 235–270.

CROMPTON, P. Extendig the Stochastic Approach to Index Numbers. *Applied Economics Letters*, 2000, 7, pp. 367–371.

DIEWERT, W. E. *On the Stochastic Approach to Index Numbers*. Discussion Paper No. 05–31, 1995, Department of Economics, University of British Columbia.

DIEWERT, W. E. *On the Stochastic Approach to Linking the Regions in the ICP*. Discussion Paper No. 04–16, 2004, Department of Economics, University of British Columbia.

DIEWERT, W. E. Weighted Country Product Dummy Variable Regressions and Index Number Formulae. *Review of Income and Wealth*, 2005, 51, pp. 561–570.

EDGEWORTH, F. Y. Measurement of Change in Value of Money I. The first Memorandum presented to the British Association for the Advancement of Science, 1887. Reprinted in: *Papers Relating to Political Economy*, Vol. 1, New York: Burt Franklin, 1925, pp.198–259.

EDGEWORTH, F. Y. Measurement of Change in Value of Money. The third Memorandum presented to the British Association for the Advancement of Science, 1889. Reprinted in: *Papers Relating to Political Economy*, New York: Burt Franklin, 1925, Vol. 1, pp. 259–297.

EDGEWORTH, F. Y. Some New Methods of Measuring Variation in General Prices. *Journal of the Royal Statistical Society*, 1888, 51, pp. 346–368.

FRISCH, R. Annual Survey of General Economic Theory: The Problem of Index Numbers. *Econometrica*, 1936, 4, pp. 1–39.

JEVONS, W. S. A Serious Fall in the Value of Gold Ascertained, and its Social Effects Set Forth, 1863. Reprinted in: FOXWELL. H. S. eds. *Investigations in Currency and Finance*. London: Macmillan, 1884.

JEVONS, W. S. The Depreciation of Gold. Letter to The Economist, 1869. Reprinted in: FOXWELL. H. S. eds. *Investigations in Currency and Finance*, London: Macmillan, 1884.

PRASADA RAO, D. S. *The Country-Product-Dummy Method: A Stochastic Approach to the Computation of Purchasig Power Parities in the ICP*. Paper presented at the SSHRC Conference on Index Numbers and Productivity Measurement, 30 June–3 July, 2004, Vancouver.

SELVANATHAN, E. A., PRASADA RAO, D. S. An Econometric Approach to the Construction of Generalised Theil-Törnqvist Indices for Multilateral Comparisons. *Journal of Econometrics*, 1992, 54, pp. 335–346.

SELVANATHAN, E. A., PRASADA RAO, D. S. *Index Numbers: A Stochastic Approach*. Ann Arbor: The University of Michigan Press, 1994.

THEIL, H. *Economics and Information Theory.* Amsterdam and Chicago: North-Holland and Rand McNally, 1967.

TÖRNQVIST, L. The Bank of Finland's consumption price index. *Bank of Finland Monthly Bulletin*, 1936, 10, pp. 1–8.

VON DER LIPPE, P. *Index Theory and Price Statistics*. Frankfurt, Germany: Peter Lang, 2007.

Progressivity of Taxes, Skeweness of Income Distribution and Violations of the Progressive Principle in Income Tax Systems

Edyta Mazurek[1] | *Wrocław University of Economics, Wrocław, Poland*

Abstract

Kakwani and Lambert state the three axioms, which should be respected by an equitable tax system. They also proposed a measurement system to evaluate the violations of the axioms. One of the axioms, axiom 2, formulates the progression principle in income tax systems. Vernizzi and Pellegrino improved the alternative index to evaluate violations concerning the progressive command in a tax system. The main aim of this paper is to compare the two indexes in order to evaluate violations of progressive principle in income tax system using the real data. We also check how the progressivity of taxes and skewness of income distribution affect the measurement of the progressive principle violation.

Keywords	**JEL code**
Personal income tax, progressive principle, redistributive effect, progressive of taxes, equitable tax system	*C81, H23, H24*

INTRODUCTION

Many authors define equity in income taxation by horizontal and vertical equity [Urban, Lambert 2008]. In this paper the equity in income taxation is defined by means of three axioms, introduced by Kakwani and Lambert in 1998. Tax system is equitable if all axioms are satisfied. Violation of them – by a personal income tax system – produces negative influence on the redistributive effect of the tax. This negative influence provides the means to characterize the type of inequity present in a tax system.

The three general rules requirement for the personal income tax system are named axioms by Kakwani and Lambert. As an axiom is defined as a mathematical statement that is accepted as being true without a mathematical proof (it is a logical statement that is assumed to be true), we propose to name

[1] Department of Statistics, Wrocław University of Economics, Komandorska 118/120, 53-345 Wrocław, Poland. E-mail: edyta.mazurek@ue.wroc.pl.

these postulates as rules. Despite their arbitrary character, tax systems that violate them intentionally are very rare. Practical solutions in personal income tax systems are not, however, so clear. Tax deduction and exemptions – commonly used tax instruments – often cause violation of these rules.

Let $x_1,x_2,...,x_n$ mean pre-tax income of n income units, who are paying $t_1,t_2,...,t_n$ in tax. We can write X as a vector of $x_1,x_2,...,x_n$ and T as a vector of $t_1,t_2,...,t_n$. In our analysis, household is set as an income unit, so:

x_1 will denote pre-tax income of household i and
t_1 tax payment of household *i*.

In this notation $y_i = x_1 - t_1$ denotes post-tax income of household *i* and $a_i = \frac{t_i}{x_i}$ - tax rate for household *i*.

The first rule – **Rule 1** – says that tax duty should increase monotonically with respect to taxpayers' ability to pay. This rule they written as:

$$x_i \geq x_j \Rightarrow t_i \geq t_j. \tag{1}$$

Because the inequalities are weak, postulate of "equal treatment of equals" could be treated as a special case of this rule. It also enables government to exempt taxpayers with the lowest incomes from having to pay tax. This rule is named *minimal progression principle*.

According to **Rule 2**, the richer people must pay taxes at higher rates. Of course, a violation of minimal progression automatically entails a violation of this principle. The weak inequalities in rule 2 mean that proportional taxation is permitted.

This second rule – *progression principle* – is defined in the following way:

$$x_i \geq x_j \text{ and } t_i \geq t_j \Rightarrow a_i \geq a_j. \tag{2}$$

If tax system is ruled out by principles 1 and 2 taken together then it means existence of regression in the tax system.

The last rule – **Rule 3** – says that a tax, which satisfies the other two rules, should cause no reranking in taxpayers' post-tax income. This rule is called no-reranking criterion and can be written as:

$$x_i \geq x_j \text{ and } t_i \geq t_j \textit{ and } a_i \geq a_j \Rightarrow x_i - t_i \geq x_j - t_j. \tag{3}$$

The Rule 3 can be seen as a vertical restriction, ruling out "too much" progression.

1 VIOLATION OF THE PROGRESSIVE PRINCIPLE

The most important for this paper is the second rule, the *progression principle* [Lambert 2001]. Violations of progression principle (also the others of rules) produces negative influence on the redistributive effect of the tax. In this context we should to be able to assess when the progression principle is not upheld and how much lost the redistributive effect produces.

The redistributive effect is defined as difference between the Gini index for pre-tax income and the Gini index for post-tax income [Lambert 2001] could be decomposed into following way [Kakwani and Lambert 1998]:

$$RE = G_X - G_{X-T} = V - S_1 - S_2 - S_3 \tag{4}$$

where: S_1 – measures loss in redistributive effect, caused by a violation of rule 1,
S_2 – loss in redistributive effect, caused by a violation of rule 2,
S_3 – loss in redistributive effect, caused by a violation of rule 3,
V – value of redistributive effect that might be achieved if all rules are upheld.

The measures in decomposition (4) are defined by Gini and concentration index.

Let $C_{Z,X}$ denotes concentration index for attribute Z. This measure is calculated in the same way as Gini index, but vector values of Z is ordered by incomes before taxation (X). If both orderings are identical (attribute Z causes no reranking of income), Gini and concentration indexes calculated for the same vector of incomes take the same value.

Whereas:

$$V = \tau \cdot \left(\left(C_{T,X} - G_X \right) + \left(G_{\frac{T}{X}} - C_{\frac{T}{X},X} \right) \right),$$

$$S_1 = \tau \cdot \left(G_T - C_{T,X} \right),$$

$$S_2 = \tau \cdot \left(\left(G_{\frac{T}{X}} - C_{\frac{T}{X},X} \right) - \left(G_T - C_{T,X} \right) \right)$$

and

$$S_3 = G_{X-T} - C_{X-T,X},$$

$$\tau = \frac{\sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} (x_i - t_i)}.$$

S_2 always takes non-negative values and according Kakwani-Lambert methodology:

The progression principle is violated

$$\Updownarrow$$

$$S_2 > 0$$

If S_2 is zero, the progression principle is upheld.

Violation of rule 1 about minimal progression automatically entails a violation of the progressive principle (rule 2). It means that income unit pairs (i,j) for which rule 1 fails cannot provide violations of the progressive principle.

In next section we check how S_2 measures violation of progression principle for real data.

2 EMPIRICAL ANALYSIS

The values of measure S_2 as a measure of violation of progression principle was analyzed on the basis of Polish data from Wrocław-Fabryczna tax office for fiscal year 2007. This set of data contains information on income and tax paid for taxpayers that file their tax return in the Municipality of Wrocław, tax office (district identification) Fabryczna. In this analysis households are equated with couples of taxpayers who take advantage of joint taxation and filled up the formulate PIT 37. The analyses were performed by author'own programmes, written in the "*R*" language.

Population of 19 487 households was divided into subpopulations with respect to the number of dependent children. We created 4 sets: family without children, family with one child, family with two children, family with three or more children.

Table 1 presents measures S_2 for each type of family and for pairs of units income which satisfied or not rule 1. When rule 1 is violated and rule 2 is upheld, the measure of loss in RE due to violation rule to should be equal 0. It is not true for S_2.

Table 1 Values of the measure S_2 for fourth type of family

Parameter	Rule 1 violated	Rule 1 upheld	Total
family without children			
S_2	0.001379	0.001717	0.003096
family with one child			
S_2	0.001213	0.000729	0.001942
family with two children			
S_2	0.001174	0.000443	0.001617
family with three or more children			
S_2	0.001582	0.00021	0.001792

Source: Own calculations

For each group of taxpayers we observe that the measures S_2 are greater than 0 for pairs of units (i,j) for which rule 1 is violated. According to Kakwani and Lambert methodology it means that, progression principle is violated. If we look at the mathematical record of the rule 2 (see formula (2)) it could be observed that for pairs of units (i,j) for which rule 1 is violated the rule 2 is not violated and the measure S_2 could be zero. If we want to use this measure S_2 we should firstly eliminate from set of data the pairs of units (i,j) for which rule 1 is violated and next calculate measure S_2. Elimination from data the pairs of units (i,j) for which rule 1 is violated is not a simple the task.

Pellegrino and Vernizzi (2013) introduced the correction of the measure of loss in redistributive effect, caused by a violation of rule 2 – S_2^* – which can be used for full set of data. The measures is defined as follows:

$$S_2^* = \frac{\tau}{2n^2} \cdot \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{a_i - a_j}{\mu_A} \cdot \left[\left(I_{i-j}^{A} - I_{i-j}^{A/X} \right) - \left(I_{i-j}^{T} - I_{i-j}^{T/X} \right) \right] \cdot p_i p_j \,, \tag{5}$$

where:

n is a sample size, $a_i = \frac{t_i}{x_i}$, p_i, p_j are weights associated to a_i and a_j, $\sum_{i=1}^{k} p_i = n$, μ_A is the average of a_i, $i=1,\ldots k$. I_{i-j}^{Z}, $I_{i-j}^{Z/X}$ are indicator function for attribute Z:

$$I_{i-j}^{Z} = \begin{cases} 1 & z_i \geq z_j \\ -1 & z_i < z_j \end{cases} \qquad I_{i-j}^{Z/X} = \begin{cases} 1 & \text{if} \quad x_i > x_j \\ -1 & \text{if} \quad x_i < x_j \\ I_{i-j}^{Z} & \text{if} \quad x_i = x_j \end{cases} .$$

Table 2 presents values of the measure S_2^* for analyzed sets of data.

Table 2 The measure S_2 for each type of family

Parameter	Rule 1 violated	Rule 1 upheld	Total
family without children			
S_2^*	0	0.001717	0.001717
family with one child			
S_2^*	0	0.000729	0.000729
family with two children			
S_2^*	0	0.000443	0.000443
family with three or more children			
S_2^*	0	0.00021	0.00021

Source: Own calculations

We can observe that if $S_2^* = 0$, it not necessarily true that S_2. The measure S_2^* is demonstrating appropriate behaviors for each of analysed data sets. In every case of pairs of units (i,j) for which rule 1 is violated the value measure S_2^* is zero correctly. It proves that S_2^* could be better measure for lost of redistributive effect due to violation of progressive principle. Table 3 presents the results of decomposition of RE according formula (4) for four Polish data sets. For families without children, the personal income tax system reduces the inequality of income by 1.6 percentage points. Losses in this redistributive effect due to violation of rule 1, 2 and 3 are 0.4 percentage points according to KL methodology or 0.3 percentage points according to VP methodology. The difference appears as a result of difference between estimation of the loss of redistributive effect due to violation of Rule 2 according to KL and VP. The inequity, resulting from violation of Rule 2, reduces overall redistributive effect by 0.31 (according to KL) percentage points which is 19.13 % of *RE* or 0.17 (according to VP) percentage points which is only 10.80 % of RE.

Table 3 RE decomposition for taxpayers divided into subpopulations with respect to the number of dependent children

family without children

	Gini for pre-tax income	Gini for post-tax income	RE	Potential equity	Rule 1	Rule 2	Rule 3	Total Rules
Kakwani and Lambert	0.371782	0.355399	0.016382	0.020589	0.000951	0.003134	0.000121	0.004206
	percentage of RE (%):		100.00	125.68	5.80	19.13	0.74	25.68
Vernizzi and Pellegrino	0.371782	0.355399	0.016382	0.019195	0.000951	0.001740	0.000121	0.002812
	percentage of RE (%):		100.00	119.19	5.92	10.80	0.75	17.46

family with 1 child

	Gini for pre-tax income	Gini for post-tax income	RE	Potential equity	Rule 1	Rule 2	Rule 3	Total Rules
Kakwani and Lambert	0.346474	0.323702	0.022772	0.025682	0.000854	0.001942	0.000115	0.00291
	percentage of RE (%):		100.00	112.78	3.75	8.53	0.50	12.78
Vernizzi and Pellegrino	0.346474	0.323702	0.022772	0.02447	0.000854	0.000729	0.000115	0.001698
	percentage of RE (%):		100.00	107.45	3.75	3.20	0.50	7.45

family with two children

	Gini for pre-tax income	Gini for post-tax income	RE	Potential equity	Rule 1	Rule 2	Rule 3	Total Rules
Kakwani and Lambert	0.346507	0.318711	0.027796	0.030329	0.000797	0.001617	0.000119	0.002533
	percentage of RE (%):		100.00	109.11	2.87	5.82	0.43	9.11
Vernizzi and Pellegrino	0.346507	0.318711	0.027796	0.029155	0.000797	0.000443	0.000119	0.001359
	percentage of RE (%):		100.00	104.89	2.87	1.59	0.43	4.89

family with three or more children

	Gini for pre-tax income	Gini for post-tax income	RE	Potential equity	Rule 1	Rule 2	Rule 3	Total Rules
Kakwani and Lambert	0.387007	0.353281	0.033726	0.036399	0.000782	0.001792	9.95E-05	0.002673
	percentage of RE (%):		100.00	107.93	2.32	5.31	0.29	7.93
Vernizzi and Pellegrino	0.387007	0.353281	0.033726	0.034817	0.000782	0.00021	9.95E-05	0.001091
	percentage of RE (%):		100.00	103.23	2.32	0.62	0.29	3.23

Source: Own calculations

It is almost twofold increase for KL methodology in comparison with VP methodology. The difference is so big that it is worth doing an investigation of why $S_2 - S_2^* > 0$and conditions when S_2 can be a reasonable approximation of S_2^*. The second aspect of this problem is the fact that value of S_2 influences on the potential redistributive effect. It is important because potential redistributive effect informs us, what is worth mentioning, how removal of inequities due to violation of rules could potentially improve the redistributive effect of taxation without increasing the marginal tax rates for the taxpayers groups.

The total inequity in the Polish tax system reduces the redistributive effect of taxation for group of family without children by 0.42 percentage points (according to KL) or by 0.28 percentage points (according to VP). These results suggest that the absence of all mentioned inequities could reduce the inequality of income by 2.06 percentage points or by 1.91 percentage points (instead of 1.64 percentage points).

For the group of taxpayers with one, two or three or children we observe similar connection between S_2.and S_2^*. Table 4 presents differences between this two values for each type of family, which are always greater than 0. The differences are from 4.22 to 8.51 points of percentage of *RE* for different type of families.

Table 4 The differences between S_2 and S_2^* for each type of family

Type of family	S_2	S_2^*	$S_2 - S_2^*$	$S_2 - S_2^*$ as percentage of RE (%)
0 children	0.003134	0.00174	0.001394	8.51
1 child	0.001942	0.000729	0.001213	5.33
2 children	0.001617	0.000443	0.001174	4.22
3 or more children	0.001792	0.00021	0.001582	4.69

Source: Own calculations

Where do the differences come from? We are looking for conditions when we can use S_2 as good approximation of S_2^*. We can give some thought if difference $S_2 - S_2^*$ depends on tax progressivity or on the skweness of income distribution. Below table presents $S_2 - S_2^*$ and the measure of tax progressivity defined by Kakwani (1977) as a difference between the concentration index of taxes and the Gini index of the pre-tax income.

$$\Pi^K = D_T - G_X. \tag{6}$$

Values of this measure are included in a range: $\Pi^K \in [-1 - G_X, 1 - G_X]$. Positive values, $\Pi^K > 0$, mean the progressive tax system. For the proportional system we receive: $\Pi^K = 0$. The negative values $\Pi^K < 0$ are describing the regressive tax system. The measure Π^K could be interpreted as the percent of total fiscal charges which remained changed from worse earning to the better earning for the effect of the progressions of tax system.

Table 5 The name and way of create data sets for different types of skwenesses

name of set	description
0 children 80%	contains 80% taxpayers with the lowest income from set 0 children
0 children 90%	contains 90% taxpayers with the lowest income from set 0 children
0 children 95%	contains 95% taxpayers with the lowest income from set 0 children
0 children 97%	contains 97% taxpayers with the lowest income from set 0 children
0 children 99%	contains 99% taxpayers with the lowest income from set 0 children
0 children 100%	taxpayers with 0 dependent children

Source: Own presentation

We can give also some thought if the difference $S_2 - S_2^*$ depends on skweness of income distribution. In order to do that we created data sets by cutting down the origin sets. In this way we created the following sets presented in Table 5.

In the same way we created the sets of taxpayers with 1, 2 or 3 dependent children. Table 6 presents results of analysis from each data sets. We calculated apart from differences $S_2 - S_2^*$, influence the differences on redistributive effect – $\frac{S_2 - S_2^*}{RE}$ as well as the skewness of income distribution and progressivity index.

Table 6 Results of analysis for created data sets

Data set	$S_2 - S_2^*$	RE	$\frac{S_2 - S_2^*}{RE}$	Skweness	Π^K
0 children 80%	0.00112	0.00693	0.1618	0.14	0.089936
0 children 90%	0.00109	0.00700	0.1562	0.44	0.086222
0 children 95%	0.00111	0.00842	0.1321	0.74	0.097939
0 children 97%	0.00114	0.00966	0.1183	0.96	0.108245
0 children 99%	0.00122	0.01202	0.1016	1.48	0.126605
0 children 100%	0.00139	0.01638	0.0851	3.82	0.157335
1 child 80%	0.000853	0.009299	0.0918	0.06	0.163209
1 child 90%	0.000871	0.011029	0.0790	0.35	0.171039
1 child 95%	0.000910	0.013396	0.0679	0.66	0.188018
1 child 97%	0.000950	0.014993	0.0634	0.90	0.198983
1 child 99%	0.001041	0.018037	0.0577	1.44	0.219137
1 child 100%	0.001212	0.022772	0.0532	3.77	0.247409
2 children 80%	0.000713	0.010824	0.0659	0.03	0.240979
2 children 90%	0.000763	0.013927	0.0548	0.38	0.252356
2 children 95%	0.000815	0.016688	0.0488	0.69	0.264156
2 children 97%	0.000856	0.018205	0.0470	0.89	0.270102
2 children 99%	0.000960	0.021927	0.0438	1.46	0.289748
2 children 100%	0.001174	0.027796	0.0422	8.23	0.317156
3 children 80%	0.000580	0.008092	0.0716	0.12	0.362515
3 children 90%	0.000738	0.012809	0.0576	0.61	0.377318
3 children 95%	0.000839	0.016909	0.0496	0.9	0.383747
3 children 97%	0.000924	0.019023	0.0486	1.23	0.384923
3 children 99%	0.001122	0.024289	0.0462	1.89	0.399589
3 children 100%	0.001582	0.033726	0.0469	6.64	0.419609

Source: Own calculations

We can observe that for the lowest tax progressivity we have the biggest value of difference $S_2 - S_2^*$ for each group. Consistently for the highest tax progressivity we have the smallest value of difference $S_2 - S_2^*$. Generally, the higher the skewness is – the higher the difference between S_2 and S_2^*. Only for *0 children 90%* set we have lowest difference $S_2 - S_2^*$ and highest skweness in compare with 0 children 80%. On the other side if we are looking for conditions when we can use S_2 as good approximation of S_2^* we should analyze influence the $S_2 - S_2^*$ on redistribution effect. 4th column in Table 6 presents this influence. We observe that the higher the skewness is, the lower the influence of the $S_2 - S_2^*$ difference on redistribution effect. This relation we observe also for *0 children 90%* set.

CONCLUSIONS

We presented and compared two measures of violations of progressivity principle: S_2 and S_2^*. We carried out an investigation for different income distribution and one tax system. We tried to understand the difference between these indexes and conditions when S_2 could be a reasonable approximation of S_2^*.

If we want to only check if progression principle is upheld or not we can use both methods: original Kakwani and Lambert or modified by Vernizzi and Pellegrino. If we want to assess the loss in the redistributive effect, caused by a violation of progression principle we should use recast index S_2^*.

We observe that the higher the skewness of income distribution is, the lower influence difference $S_2 - S_2^*$ on the redistributive effect. We observe similar simple correlation between the tax progressivity index and influence difference on the redistribution effect.

References

KAKWANI, N. C. Measurement of tax progressivity: an international comparison. *Economic Journal of Political Economy*, 1977, 87, pp. 71–80.

KAKWANI, N. C., LAMBERT, P. J. On measuring inequity in taxation: A new approach. *European Journal of Political Economy*, 1998, Vol. 143, pp. 369–380.

LAMBERT, P. J. *The Distribution and Redistribution of Income.* Manchester and New York: Manchester University Press, 2001.

LAMBERT, P. J., ARONSON, J. R. Decomposition Analysis and the Gini Coefficient Revisited. *The Economic Journal*, 1993, Vol. 103, No. 420, pp. 1221–1227.

PELLEGRINO, S., VERNIZZI, A. On Measuring Violations of the Progressive Principle in Income Tax Systems. *Empirical Economics*, 2013, 45, pp. 239–245.

URBAN, I., LAMBERT, P. J. Redistribution, horizontal inequity and reranking: how to measure them properly. *Public Finance Review,* 2008, 36(5), pp. 563–587.

Application of Robust Regression and Bootstrap in Productivity Analysis of GERD Variable in EU27

Dagmar Blatná[1] | *University of Economics, Prague, Czech Republic*

Abstract

The GERD is one of Europe 2020 headline indicators being tracked within the Europe 2020 strategy. The headline indicator is the 3% target for the GERD to be reached within the EU by 2020. Eurostat defines "GERD" as total gross domestic expenditure on research and experimental development in a percentage of GDP. GERD depends on numerous factors of a general economic background, namely of employment, innovation and research, science and technology. The values of these indicators vary among the European countries, and consequently the occurrence of outliers can be anticipated in corresponding analyses. In such a case, a classical statistical approach – the least squares method – can be highly unreliable, the robust regression methods representing an acceptable and useful tool. The aim of the present paper is to demonstrate the advantages of robust regression and applicability of the bootstrap approach in regression based on both classical and robust methods.

Keywords	***JEL code***
LTS regression, MM regression, outliers, leverage points, bootstrap, GERD	*C19, C49, O11, C13*

INTRODUCTION

GERD represents total gross domestic expenditure on research and experimental development (R&D) as a percentage of GDP (Eurostat), R&D expenditure capacity being regarded as an important factor of the economic growth. GERD is one of Europe 2020 indicator sets used by the European Commission to monitor headline strategy targets for the next decade –A Strategy for Smart, Sustainable and Inclusive Growth (every country should invest 3% of GDP in R&D by 2020). GERD comprises expenditure of four institutional sectors of production – business enterprise, government, higher education and private non-profit establishments. Expenditure data involve the research funds allocated in the national territory, regardless of their source.

Generally GERD depends on various elements of a general economic background, such as the employment, innovation, research, science and technology. Both GERD and the above indicators' values

[1] University of Economics, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. Email: blatna@vse.cz.

vary among the European countries and, consequently, the occurrence of outliers can be envisaged in the EU countries' GERD analysis.

A classical statistical approach to regression analysis – the least squares method (LS) – can be highly unsatisfactory due to the presence of outliers that are likely to occur in an analysis of any real data. In such a case, robust regression becomes an acceptable and useful tool, since it provides a good fit to the bulk of the data, the outliers being exposed clearly enough. The aim of this paper is to verify the applicability of the robust regression and bootstrapping (resampling) technique based on both LS and robust regression, the economic GERD analysis not being its main objective.

1 LITERATURE

Robust regression techniques are rarely used in economic analysis; only a few applications can be found in the available literature. Zaman, Rousseeuw, Orhan (2001), for instance, applied a high breakdown robust regression method to three linear models, having compared regression statistics for both the LS technique used in the original paper and the robust method. The authors eventually recommended that robust techniques should be used to avoid the confusion effect of "bad" leverage points leading to a significant bias of the regression results. Finger, Hediger (2007) promoted the application of robust instead of LS regression for the estimation of agricultural and environmental production function and Colombier (2009) also estimated the growth effects of OECD fiscal policies having employed robust methods.

Numerous analyses of R&D expenditure have been made on the basis of different criteria such as the source of funds, field of science, type of costs, economic activity, enterprise size class, socioeconomic objectives, regions, etc. Guellec (1997, 2001) dealt with the cause of fluctuations in investments in R&D and the connection between GERD and productivity growth. Kroll, Zenker (2009) looked into the development of R&D expenditure at a regional level and Zhang (2006) published the results of an empirical analysis of national energy R&D expenditures. Since the launch of Europe 2020 strategy, a lot of studies, papers and reports have been released. Commenting on the strategy, some of them make relevant remarks regarding the 3% target for the GERD indicator to be reached within the EU by 2020. Albu (2011), for example, investigated to what extent the EU members complied with the R&D investment targets set by Europe 2020 strategy, their actual spending being below 2% of GDP on average and only three member states reporting the R&D expenditure ratio to be higher than 3% of GDP. Dachs (2012) analyzed an economic impact of the internationalization of business investments in R&D, Spišáková (2013) examinined the influence of the economic crisis on the achievement of Europe 2020 target in the R&D area.

2 METHODOLOGY

2.1 The principle of robust regression

Robust regression techniques are an important complement to the classical least squares (LS) regression method. Robust techniques produce results similar to LS regression when the data are linear with normally distributed errors. The results, however, can differ significantly when the errors do not satisfy normality conditions or when the data contain outliers. Robust regression is an alternative to LS regression when the data are contaminated with outliers or influential observations. It can be used for detecting influential observations as well.

It is a common practice to distinguish between two types of outlying observations in the regression, those in the response variable representing a model failure. Such observations are called outliers in the y-direction or vertical outliers, those with respect to the predictors being labelled as leverage points. The leverage point is defined as $(x_{k_1},..., x_{k_p}, y_k)$ for which $(x_{k_1},..., x_{k_p})$ is outlying with respect to $(x_{i_1},..., x_{i_p})$ in the data set. Regression outliers (influential points) are the cases for which $(x_{k_1},..., x_{k_p}, y_k)$ deviates from the linear relation followed by the majority of the data, both the explanatory and response variable being taken into account simultaneously.

First, let us briefly mention the principles of selected robust methods used in our analysis. In robust regression, an important role is played by the breakdown point which is the fraction of “bad” data that the estimator can tolerate without being affected to an arbitrarily large extent. Having a zero breakdown value, even a small proportion of deviant observations can cause systematic distortions in LS regression estimates. Two regression methods with a high breakdown point were employed. The least trimmed squares (LTS) estimator (proposed by Rousseeuw 1984)) is obtained by minimizing $\sum_{i=1}^{h} r_{(i)}^2$, where $r_{(i)}^2$ the i-th order statistic among the squared residuals written in the ascending order, h is the largest integer between $[n/2]+1$ and $([n/2]+[(p+1/2)])$, p is the number of predictors (including an intercept) and n is the number of observations. The usual choice $h \approx 0.75n$ yields the breakdown point of 25 % - see Hubert, Rousseeuw, Van Aelst (2008).

LTS regression with a high breakdown point is a reliable data analytic tool that can be used to detect vertical outliers, leverage and influential points (observations whose inclusion or exclusion result in substantial changes in the fitted model) in both simple and multivariate settings. A more detailed description is available in, e.g., Ruppert, Carroll (1980), Rousseeuw (2003), Chen (2002), Fox (2002) or Hubert, Rousseeuw, Van Aelst (2008).

MM-estimates (proposed by Yohai (1987) combine a high breakdown point with good efficiency (approximately 95% to LS under the Gauss-Markov assumption). MM regression is defined by a three-stage procedure (for details, see Yohai (1987), Chen (2002) or Rousseeuw (2003)). At the first stage, an initial regression estimate is computed; it is consistent, robust, with a high breakdown point but not necessarily efficient. At the second stage, an M-estimate of the error scale is computed, using residuals based on the initial estimate. Finally, at the third stage, an M-estimate of the regression parameters based on a proper redescending ψ-function is computed by means of the formula:

$$\sum_{i=1}^{n} \mathbf{x}_i \psi\left(\frac{y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right) = 0, \tag{1}$$

where $\hat{\sigma}$ stands for a robust estimation of the residual standard deviation (calculated in the 2nd step) and $\psi = \rho'$ is the derivation of the proper loss function ρ. A more detailed description of robust regression methods is available in Chen (2002), Rousseeuw (2003), Fox (2002), Yohai (1987), SAS and SPLUS manuals. Due to SAS and S-PLUS software used in the analysis, Tukey’s bisquare loss function was employed:

$$\rho(e) = \begin{cases} \dfrac{k^2}{6}\left\{1 - \left[1 - \left(\dfrac{e}{k}\right)^2\right]^3\right\} & \text{for } |e| \le k \\ \dfrac{k^2}{6} & \text{for } |e| > k \end{cases}, \tag{2}$$

where e means residuum, the tuning constant k = 4.685 for the bisquare loss function.

2.1.1 Identification of outliers, leverage and influential points

Extensive numerical and graphical diagnostic methods for detecting outliers and influential observations can be used. For more details, see, e.g. Rousseeuw, Van Zomeren (1990), Rousseeuw (2003), Fox (2002), Olive (2002), Chen (2002). In this paper, the following methods have been employed:

- *Residuals associated with LTS regression;*
- *Standardized residuals* (the residuals divided by the estimates of their standard errors, the mean and standard deviation equalling 0 and 1 respectively);
- *Studentized residuals* (a type of standardized residuals follows at t distribution with n-p-2 Df), attention being paid to studentized residuals that exceed ± 2.5 (or ± 2.0);

- *The robust distance* defined as:

$$RD(x_i) = \sqrt{[x_i - \mathbf{T}(\mathbf{X})]^T \, \mathbf{C}(\mathbf{X})^{-1}[x_i - \mathbf{T}(\mathbf{X})]}\,, \tag{3}$$

where $\mathbf{T}(\mathbf{X})$ is the robust location estimates vector and $\mathbf{C}(\mathbf{X})$ is the scatter matrix for the matrix of covariates;

- *Diagnostic plots* provided as fundamental data mining graphical tools for quick identification of an outlier, determine whether outliers have influence on classical estimates. In order to visualize vertical outliers and leverage points, the following plots were used:
  - *regression diagnostic plot* (a plot of standardized residuals of robust regression versus robust distances $RD(x_i,)$),
  - *plot of standardized residuals versus their index,*
  - *normal Q-Q plot of standardized residuals* and
  - *plot of kernel estimate of residuals´ density.*

2.2 The principle of bootstrap in regression

The bootstrap was introduced by Efron (1979). Bootstrapping is a general approach to statistical inference based on replacement of the true sampling distribution for a statistic by resampling from the original observed data (the original sample of size n). Bootstrap technique assumes only finite values of some moments, but hardly any restricting assumptions about the underlying probability distribution. It replaced classical methods' assumptions with complex calculations for the correctness assessment of a relationship found within a particular sample. The fundamental element of bootstrap is a bootstrap sample. The resampling procedure in regression brings R artificial samples of n pairs of observations from the data in the original observed sample. For bootstrapping pairs in regression models, the bootstrap sample is selected by simple random sampling observations (i.e. the response value and the corresponding vector of independent regressor variables) without replacement. Then standard errors, confidence intervals and the bias of bootstrap parameter estimates are calculated. The bias is estimated by the difference between an average bootstrapped value of the regression coefficient and its original-sample value. The bootstrap percentile interval (EP) is based on empirical quantiles of the bootstrap regression coefficients b_b^*, while the bias-corrected, accelerated percentile interval (BC_a) with correction factors for lower and upper percentiles is grounded on the jackknife values of the statistic β (see, e.g. Cole (1999), DiCiccio, Efron (1996), Freedman (1981), Efron (1993, 2000), Stine (1990). The resampling distribution of the regression coefficients is then constructed empirically by resampling from the sample.

In the bootstrap regression procedure, the least squares (LS) method is often used to estimate the parameters of regression models. It is, however, extremely sensitive to outliers and non-normality of errors. The robust bootstrapping method replaces the classical bootstrap mean and standard deviation with robust estimates, using robust regression estimates with a high breakdown point. In our analysis, MM-regression with initial LTS estimates has been used. The bootstrap is not used for regression parameters estimation, being a tool for the acquisition of confidential intervals and bias regression parameters estimation.

3 RESULTS AND DISCUSSION

The following regression methods have been employed in an analysis of the GERD in EU27 countries:

- least squares regression (LS),
- least trimmed squares regression (LTS),
- MM-regression (MM),
- bootstrap regression based on the LS method (B),
- bootstrap regresion based on robust MM-regression (RB).

The analysis is based on 2010 data, calculations being performed by means of SAS 9.2 and S-Plus 6.2 statistical software. All the data as well as indicator definitions have been adopted from the Eurostat database.[2] The economic indicators employed in the analysis are given in the appendix to this paper.

The GERD (total gross domestic expenditure on research and experimental development as a percentage of GDP) is one of Europe 2020 headline indicators being tracked within the Europe 2020 strategy. The headline indicator is the 3 % target for the GERD to be reached within the EU by 2020. "This target has succeeded in focusing attention on the need for the both the public and private sectors to invest in R&D but it focuses on input rather than impact" (see European Commision, 2010, p. 8). From this point of view, GERD is consider as a dependent variable in the analysis.

For the GERD as the dependent variable, numerous linear regression models have been tested using the least squares linear regression (LS) and robust MM-regression. Identification of vertical outliers, leverage points and influential points was performed using LTS regression. SAS uses the default value $h = [(3n + p + 1)/4]$. For $n = 27$ and $p = 3$ or $p = 4$, we get $h = 21$, and the corresponding breakdown point of about 21–25%. The existence of vertical outliers or leverage points in the model can be quickly identified from the robust diagnostic plot, LS diagnostics being on the left and robust diagnostics on the right side. Horizontal broken lines are located at +2.5 and –2.5 and the vertical line is located at the cutoffs of $\pm\sqrt{\chi^2_{p-1;0.975}}$, where p is the number of predictors. The points lying to the right of the vertical line are leverage points, those lying above or below horizontal lines are regarded as vertical outliers. In the case of classical LS regression, the classical index of determination (R-squared) and the results of significance t-tests and F-tests (at a significance level of 5%) were used. In the case of robust regression, the decision which of the alternative models should be preferred was based on robust diagnostic selection criteria: the robust index of determination (R-squared), significance robust Wald and F-test and robust selection information criteria –Robust Akaike's Information Criterion (*AICR*), Robust Bayesian Information Criterion (*BICR*) and Robust Final Prediction Error (*RFPE*), (see e.g. Hampel (1983), Hampel, Ronchetti, Rousseeuw, Stahel (1996), Ronchetti (1985), Sommer, Huggins (1996), SAS and SPLUS manuals). In both LS and robust regressions, the normality of residuals was also taken into consideration to determine which model ought to be preferred. Numerous regression models, using the set of indicators (predictors) available from the Eurostat database, have been computed. For regression models that fulfill the aforementioned criteria, both classical and robust bootstrapping regression were applied as well. In the analysis, only models with two or three regressors were fully acceptable. The selected models– mutually different from the statistical point of view – are presented, the occurrence and variety of outliers being crucial for their choice. In all tables, t denotes the test statistic related to individual t-tests, p-value expresses the minimal significance level, where the null hypothesis can be rejected, R-sq. denoting the index of determination.

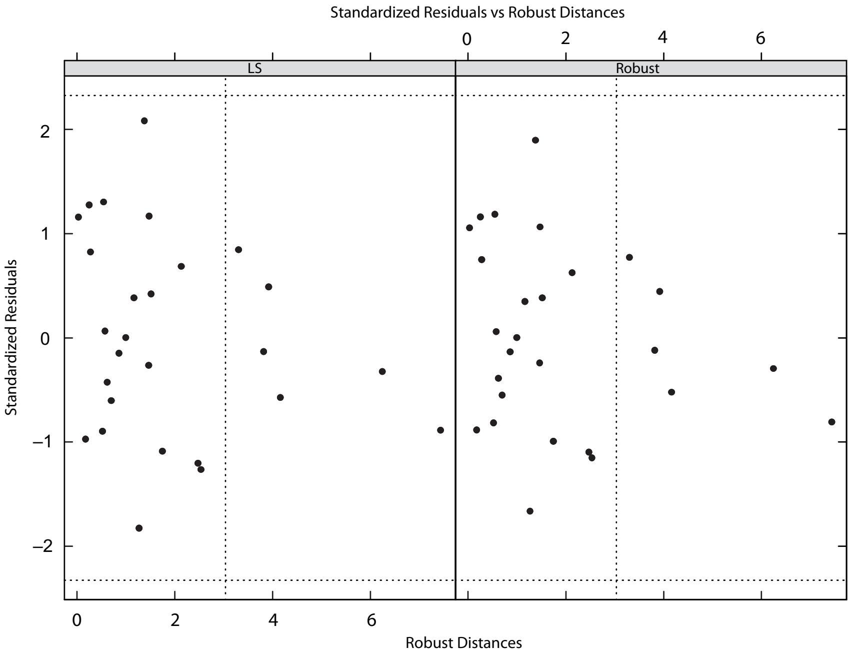
In the presented models the following predictors have been included:

CPL Comparative Price Level (EU27 = 100%);
ER Employment rate total (the ratio of employed persons aged 20–64 and the total population of the same age group;
HICP Harmonised indices of consumer prices (2005 = 100);
IRUI Individuals regularly using the Internet (in percent; frequency of Internet access: once a week);
LPH Labour productivity per hour worked;

In the first model that includes explanatory variables CPL and IRUI, both LS and robust diagnostics identified six leverage points, none of them, however, being also an vertical outlier (see Figure l).

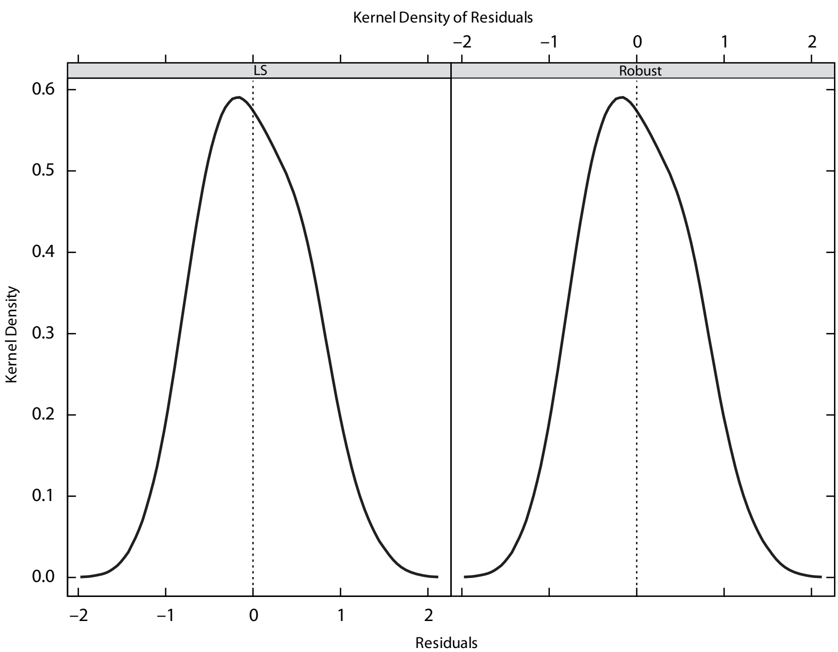
[2] <http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database> and / or <http://apl.czso.cz/pll/eutab/html>.

Figure1 Diagnostic Plot (GERD~CPL+IRUI model)



Source: Author's own elaborations

Figure 2 Kernel estimate of residuals' density (GERD~CPL+IRUI model)



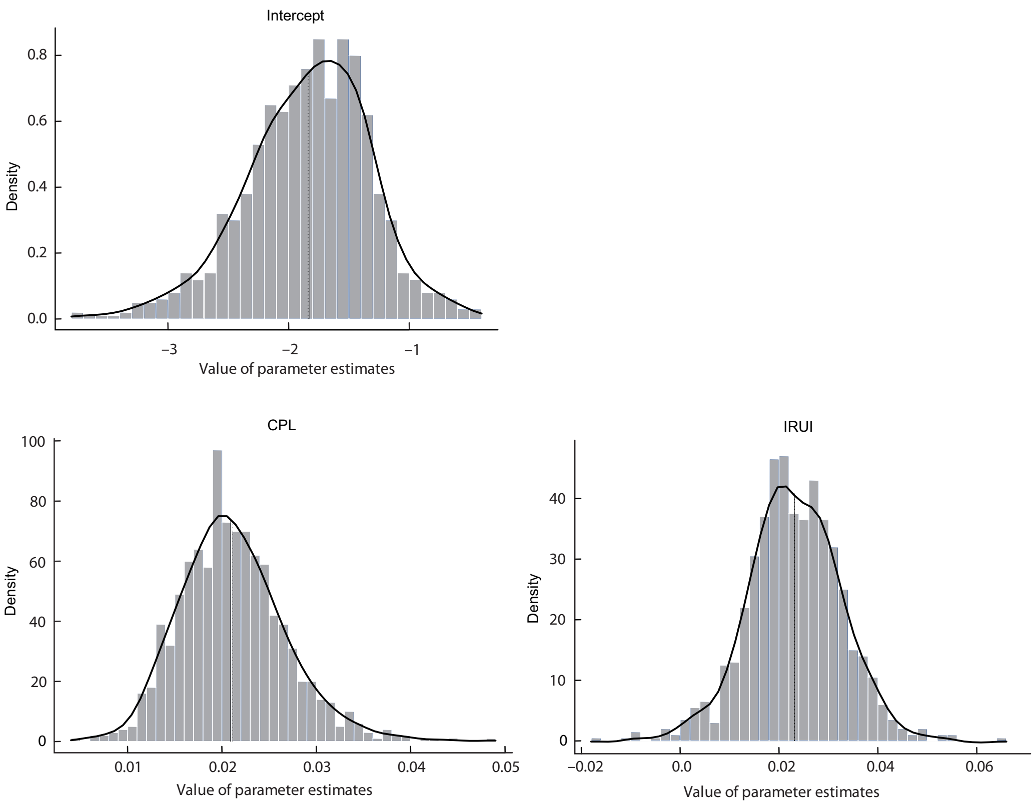
Source: Author's own elaborations

Table 1 Classical and robust bootstrap regression, LS and MM regression for GERD ~ CPL + IRUI model

		Observed	Bias	Mean	SE	95% EP	95% BCa
B R = 1000	Interc.	–1.8247	–0.0003	–1.8250	0.5101	–2.909; –0.906	–3.0039; –0.979
	CPL	0.0209	0.0004	0.0213	0.0053	0.0112; 0.032	0.0112; 0.032
	IRUI	0.0231	–0.0005	0.0226	0.0096	0.003; 0.042	0.0035; 0.0423
RB R = 1000	Interc.	–1.8247	–0.0003	–1.8250	0.5101	–2.909; –0.906	–3.0040; –0.979
	CPL	0.0209	0.0004	0.0213	0.0053	0.0112; 0.032	0.0112; 0.0321
	IRUI	0.0231	–0.0005	0.0226	0.0096	0.003; 0.042	0.0035; 0.0422
		Parameter	**SE**	**t**	**p-value**	**95% conf. interval**	
LS R-sq. 0.6629	Interc.	–1.8247	0.5175	–3.526	0.0017	–2.8928; –0.7566	
	CPL	0.0209	0.0065	3.2344	0.0035	0.0076; 0.0343	
	IRUI	0.0231	0.0099	2.3347	0.0283	0.0027; 0.0435	
MM R-sq. 0.5724	Interc.	–1.8247	0.6539	–2.790	0.0102	–2.8391; –0.7566	
	CPL	0.0209	0.0081	2.5927	0.0160	0.0082; 0.0336	
	IRUI	0.0231	0.0123	1.8834	0.0718	0.0037; 0.0425	

Source: Data EUROSTAT, author's own calculations

Figure 3 Histograms for classical replications of regression coefficients for GERD ~ CPL + IRUI model (R = 1 000)



Source: Author's own elaborations

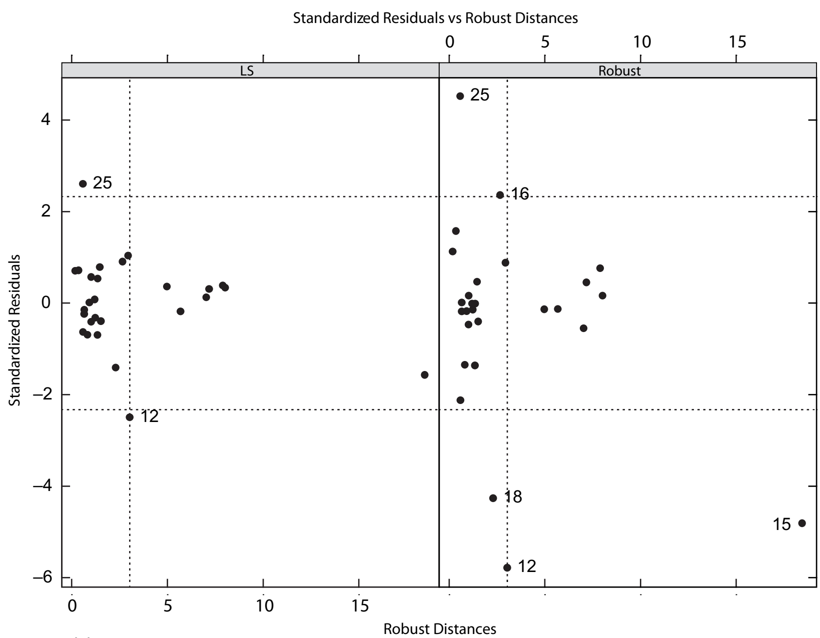
Since no vertical outliers are identified, the LS and MM-regression models are identical (see Table 1), classical and robust bootstraps provide the results very close to the values of the estimated regression coefficients of LS and MM-regression fits. Kernel estimates of residuals' density are almost normal but are not centred around zero both for LS and MM regression models (see Figure 2). Classical bootstrap provides the lowest standard errors and the narrowest confidence intervals of the estimated regression coefficients; they are even narrower than LS ones (for any regression coefficients). The bias is a difference between an average bootstrapped value of the regression coefficient and its original sample value. Histograms of regression coefficients' estimates are adequately symmetric in both bootstrap methods, robust bootstrap, however, providing broader confidence intervals. Histograms of regression coefficients' estimates for classical bootstrap see in Figure 3.

Due to the absence of vertical outliers, both classical regression and classical bootstrap are fully appropriate in the model with explanatory CPL and IRUI variables. The dependence can be expressed in the form:

$$\text{GERD} = -1.8247 + 0.0209\ \text{CPL} + 0.0231\ \text{IRUI}. \tag{4}$$

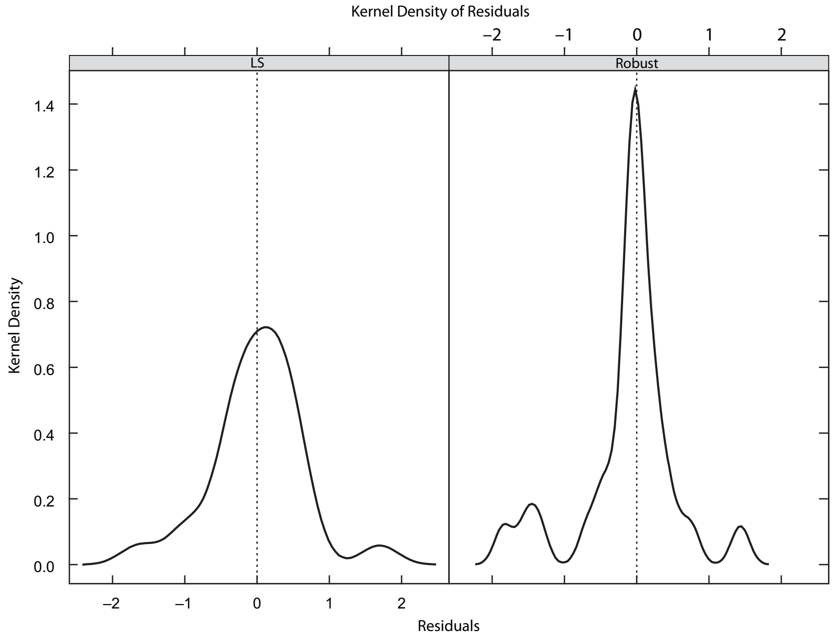
The index of determination R-sq. equals 0.6629. Both the explanatory variables have a positive influence on GERD, the partial coefficients being statistically significant at a 3% level at least. Comparative price levels (CPL) indicie the ratio between purchasing power parities (PPPs) and the market exchange rate in a particular country. The ratio is calculated in relation to the EU average (EU27 = 100). If the CPL index for a country is higher/lower than 100, the country concerned is relative expensive/cheap compared to the EU average. CPL is a measure of a nominal convergence. IRUI expresses the percentage of individuals regularly using the internet; it is one of indicators of information society expressing computer literacy of a country. In the EU countries, both a higher CPL value and a higher computer literacy, are connected with a higher expenditure on R&D. This conclusion is in general conformity with the European Commission recommendations in the area of "smart growth" promotion in the EU.

Figure 4 Diagnostic Plot (GERD ~ ER + LPH model)



Source: Author's own elaborations

Figure 5 Kernel estimate of residuals' density (GERD ~ ER + LPH model)



Source: Author's own elaborations

Table 2 Classical and robust bootstrap regression, LS and MM regression for GERD ~ ER + LPH model

		Observed	Bias	Mean	SE	95% EP	95% BCa
B R = 1000	Interc.	–5.4990	–0.0645	–5.5638	1.8579	–9.355; –1.974	–9.0248; –1.5642
	ER	0.0915	0.0001	0.0917	0.0309	0.0318; 0.153	0.0282; 0.1514
	LPH	0.0090	0.0007	0.0097	0.0045	0.0020; 0.019	–0.0000; 0.018
RB R = 1000	Interc.	–7.0419	0.0366	–7.005	2.0862	–10.98; –1.422	–8.8610; 1.6918
	ER	0.1108	0.0001	0.1109	0.0322	0.0215; 0.172	–0.0229; 0.1383
	LPH	0.0126	–0.0008	0.0118	0.0044	0.0007; 0.016	–0.0011; 0.0159
		Parameter	**SE**	**t**	**p-value**	**95% conf. interval**	
LS R-sq. 0.5639	Interc.	–5.4990	1.6750	–3.2830	0.0031	–8.9560; –2.0421	
	ER	0.0916	0.0265	3.4565	0.0021	0.0367; 0.1463	
	LPH	0.0090	0.0041	2.2213	0.0360	0.0006; 0.0173	
MM R-sq 0.5380	Interc.	–7.0419	1.6958	–4.1525	0.0004	–8.5971; –3.7757	
	ER	0.1108	0.0271	4.0953	0.0004	0.0664; 0.1435	
	LPH	0.0127	0.0045	2.8188	0.0095	0.0009; 0.0122	

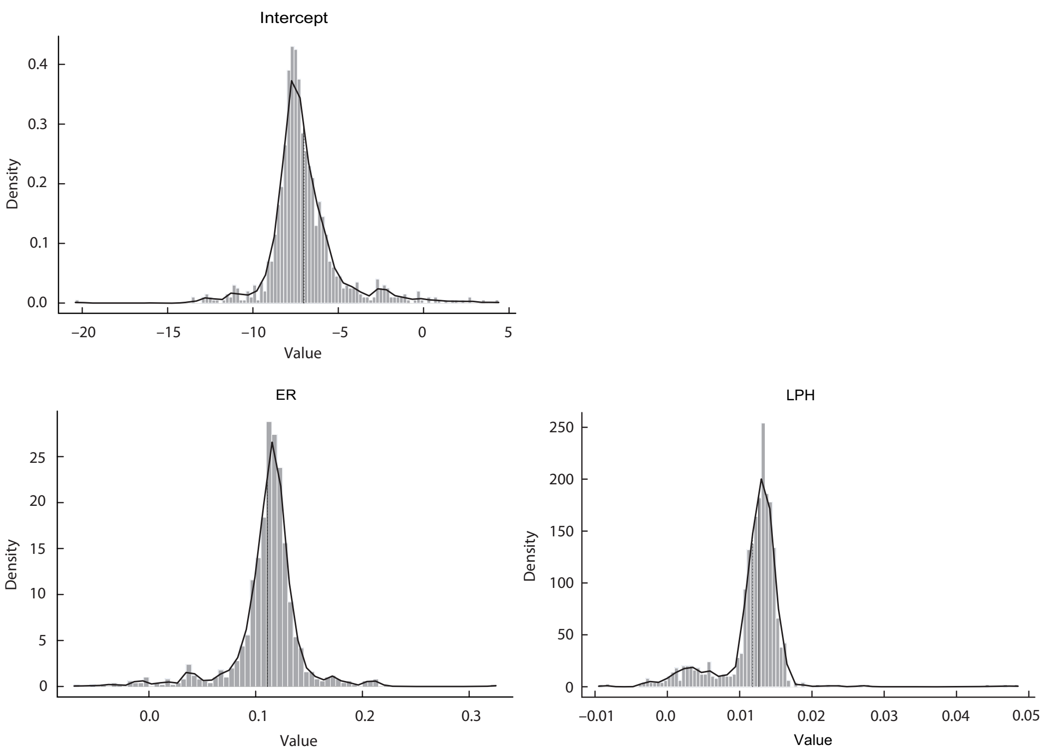
Goodness-of-fit tests for robust MM model	AICR	BICR	RFPE
	22.53	29.958	24.258

Source: Data EUROSTAT, author's own calculations

The last model includes exploratory variables ER and LPH. This model is quite distinct from the previous ones. Robust diagnostics reveal four vertical outliers (12 Cyprus, 15 Luxembourg, 18 Netherlands, 25 Finland) and seven leverage points. Two observations (12 Cyprus, 15 Luxembourg) are vertical outliers and leverage points simultaneously. These observations are thus identified as influential points. Classical diagnostics reveal only two vertical outliers and seven leverage points, none of them being identified as an influential point (see Figure 4). In such a case, the differences between classical and robust models are anticipated. For fitted values, see Table 2.

Multimodality of the kernel estimate of residuals' density plot (see Figure 5) confirms the presence of outlier points. The same is apparent from histograms of the regression coefficient estimates obtained by robust bootstrapping (Figure 6). Robust bootstrap provides tightly concentrated and markedly heavy-tailed distributions as a consequence of the existence of outliers. Robust bootstrap can be used as well, despite providing slightly biased estimates. It has to be taken into account, however, that the regression coefficients have higher standard errors and wider confidence intervals than those in the MM model (see Table 2).

Figure 6 Histograms for robust replications of regression coefficients for GERD – ER + LPH model (R = 1000)



Source: Author's own elaborations

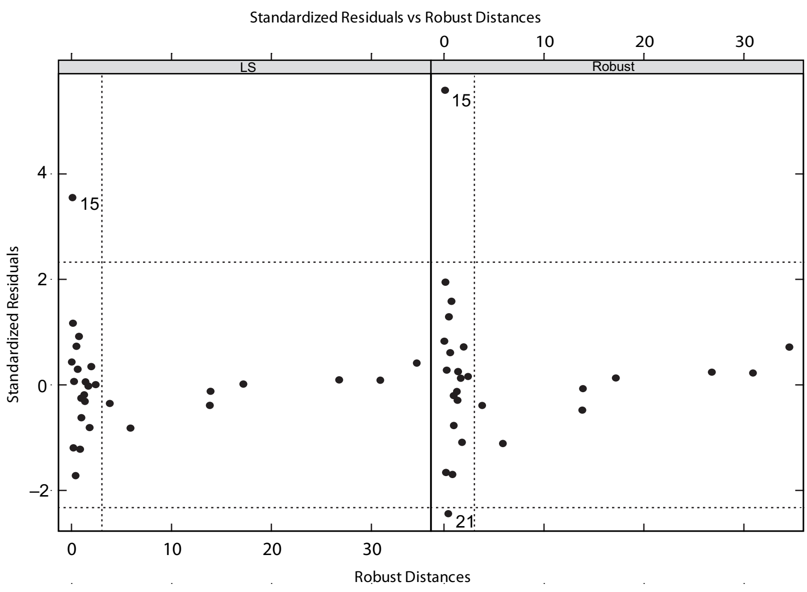
Due to the existence of influential points, the model estimated by robust regression has to be preferred. It is obvious that improper use of the classic LS regression model with significant variables without adequate identifications of outliers and testing of the normality of residuals, can lead to the acceptance of an incorrect LS model.

The exploratory variable ER (employment rate) is an indicator of labour market conditions. An increasing employment rate can lead to a decline in the percentage of GDP destined for unemployment and social security benefits, thus creating prerequisites for an increase in the proportion of GDP spent on research and development. LPH (labour productivity per hour worked) is intended to give a picture of the produktivity of national economies expressed in relation to the European Union average. If the index of a country is higher than 100, this country's level of GDP per hour worked is higher than the EU average. LPH is then a measure for the economic activity. The high level of economic activity and better working conditions are prerequisites for increasing the ratio of R&D expenditure. This could be expressed by the robust model:

$$\text{GERD} = -7.0419 + 0.1108\ \text{ER} + 0.0127\ \text{LPH}. \qquad (5)$$

In the economic literature, the GERD indicator is more frequently perceived as a factor of labour productivity growth. In the analysed period (2010), the value of the Pearson correlation coefficient between GERD and LPH was 0.5888, the value of the robust correlation coefficient being 0.4744. We presented one of suitable regression model with regressors GERD and HICP (harmonised indices of consumer prices). In this model, both LS and robust diagnostics reveal the same vertical outlier (15 Luxembourg) and seven leverage points (see Figure7). Robust diagnostics identify another vertical outlier (21 Portugal). None of them is an influential point. Multimodality of the robust regression kernel estimate of residuals' density (see Figure 8) validates the presence of outlier points.

Figure 7 Diagnostic Plot LPH ~ GERD + HICP model

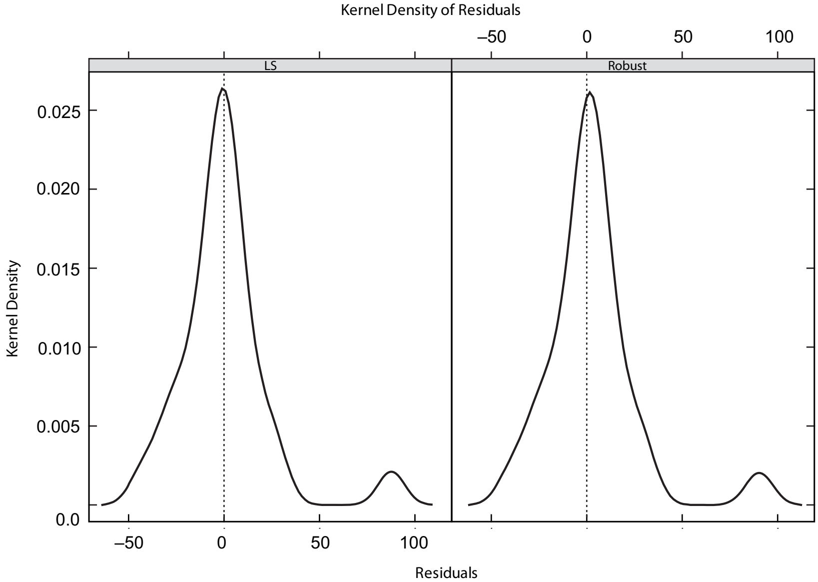


Source: Author's own elaborations

For the results of fits see Table 3. As far as GERD and HICP regressions with LPH as a dependent variable, the regression coefficients of both regressors are statistically significant (at a 5% level). The statistically significant regression coefficients indicate a positive influence of the ratio of R&D expenditure and a negative influence of inflation on labour productivity per hour worked. The resulting model estimated by robust regression has a form of:

$$\text{LPH} = 305.3371 + 11.4531\ \text{GERD} - 2.0088\ \text{HICP}. \qquad (6)$$

Figure 8 Kernel estimate of residuals' density LPH ~ GERD + HICP model



Source: Author's own elaborations

Table 3 Classical and robust bootstrap regression, LS and MM regression for LPH ~ GERD + HICP model

		Observed	Bias	Mean	SE	95% EP	95% BCa
B R = 1000	Interc.	315.3248	3.7909	319.116	54.4330	224.47; 448.696	224.442; 448.661
	GERD	11.2297	0.4133	11.643	4.3786	3.0265; 20.3203	0.4397; 19.0223
	HICP	–2.0699	–0.0349	–2.105	0.4004	–3.0265; –1.4023	–2.9392; –1.3644
RB R = 1000	Interc.	305.3371	26.3859	331.723	190.089	71.232; 808.863	50.293; 707660
	GERD	11.4531	0.7014	12.154	9.411	–5.5374; 35.064	–3.7961; 37.1874
	HICP	–2.0088	–0.2294	–2.238	1.643	–6.5055; –0.2633	–5.6710; –0.1185
		Parameter	**SE**	**t**	**p-value**	**95% conf. interval**	
LS R-sq. 0.5605	Interc.	315.3248	76.1644	4.1401	0.0004	158.1292; 472.5204	
	GERD	11.2297	6.0305	1.8622	0.0749	–1.2166; 23.676	
	HICP	–2.0699	–0.6059	–3.4164	0.0023	–3.3204; –0.8195	
MM R-sq 0.6117	Interc.	305.3371	69.4729	4.3951	0.0002	197.0488; 406.8435	
	GERD	11.4531	5,4362	2.1068	0.0458	2.8750; 19.3544	
	HICP	–2.0088	0.5514	–3.6429	0.0023	–2.8062; –1.1466	
Goodness-of-fit tests for robust MM model					AICR	BICR	RFPE
					22.2319	29.1046	18.3399

Source: Data EUROSTAT, author's own calculations

CONCLUSIONS

The GERD (total gross domestic expenditure on research and experimental development as a percentage of GDP) is one of Europe 2020 headline indicators being tracked within the Europe 2020 strategy. The headline indicator is the 3% target for the GERD to be reached within the EU by 2020.

GERD is composed of expenditure of four institutional sectors of production (business enterprise, government, higher education and private non-profit organizations). The EU countries are distinct in their structure of GERD and the ways of increasing the ratio of R&D expenditure, depending on their economic policies. In general, the value of GERD is closely linked with the country's economic development, labour market conditions and computer literacy of the population. The economic GERD analysis, however, was not the main focus of the present paper.

The statistical conclusions are not based exclusively on the results produced in this paper, but also on economic theories and research findings of the GERD variable analysis that are not explicitly referred to.

When the vertical outliers are not identified in the data, errors being normally distributed, classical LS regression is a fully appropriate method and should be preferred. In such a case, classical bootstrap regression provides even more accurate estimates of the regression parameters (with smaller standard errors and narrower confidence intervals) than LS regression. Classical bootstrap outstrips robust methods in all cases when the vertical outliers are not identified and errors are normally distributed regardless of the existence of leverage points. This conclusion was demonstrated in the GERD ~ CPL+ IRUI model.

In models with detected vertical outliers, robust regression ought to be preferred since it produces the best results. Problems with the outliers in bootstrap regression can be resolved using robust bootstrap methods. Robust bootstrap in such cases gives results similar to robust regression, but the confidence intervals are wider than the robust regression ones. This conclusion is relevant when the outliers in both x-direction (leverage points) and in y-direction (vertical outliers) are detected. With an increasing outlier's proportion, the accuracy of bootstrap estimates of the regression parameters declines. This conclusion is observed in LPH ~ GERD + HICP model.

In cases where more vertical outliers and leverage points are detected, robust regression should be preferred. The bootstrap distribution may be a rather poor estimator of the regression estimates' distribution. These results are relevant for both classical and robust bootstrap because of the proportion of the outliers in bootstrap samples which can be higher than that in the original dataset. Outlying and non-outlying observations have the same chance of belonging to any bootstrap sample and, consequently the proportion of outliers in a bootstrap sample can be even larger than the fraction of outliers that can be tolerated by robust estimates. Thus the distributions of the regression parameters have heavy tails, the confidence intervals of the regression parameters being wide. This conclusion is manifested by the results of the GERD ~ ER + LPH model.

To sum up, the findings of this study indicate that in situations when the vertical outliers are identified, robust regression with a high breakdown point ought to be given preference. It is evident that improper use of the classical LS regression model with significant variables without corresponding identifications of outliers and assessment of residual normality can lead to the acceptance of an incorrect LS model.

References

ALBU, N. Resarch and Development spending in the EU: 2020 growth strategy in perspective. *SWP Working Paper FG 1. SWP,* Berlin, 2011/Nr. 08.

CHEN, C. Robust Regression and Outlier Detection with the ROBUSTREG procedure [online]. *SUGI Paper,* SAS Institute Inc., Cary, NC., 2002. <http://www2.sas.com/proceedings/sugi27/p265-27.pdf>.

COLE, S. R. Simple bootstrap statistical inference using the SAS system. *Computer Methods and Programs in Biomedicine*, 1999, 60, pp. 79–82.

COLOMBIER, C. Growth Effects of Fiscal Policies: An Application of Robust Modified M-Estimator. *Applied Economics,* Vol. 41, Issue 7, 2009, pp. 899–912.

DACHS, B., KAMPIK, F., SCHERNGELL, T., ZAHRADNIK, G., HANZL-WEISS, D., HUNYA, G., FOSTER, N., LEITNER, S., STEHRER, R.,URBAN, W. *Internationalisation of business investments in R&D and analysis of their economic impact.* Luxembourg: Publications Office of the European Union, 2012.
DICICCIO, T. J., EFRON, B. Bootstrap confidence intervals. *Statistical Science,* 1996, 11(3), pp. 189–212.
EFRON, B. Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 1979, 7, pp. 1–26.
EFRON, B. The bootstrap and Modern Statistics. *JASA,* 2000, 95(452), pp. 1293–1296.
EFRON, B., TIBSHIRANI, R. J., eds. *An Introduction to the Bootstrap.* New York: Chapman&Hall/CRC, 1993.
EUROPEAN COMMISSION. *EUROPE 2020. A European Strategy for smart, sustainable and inclusive grows* [online]. Brussels, COM, 2010. <http://ec.europe.eu/eu2020/pdf>.
FINGER, R., HEDIGER, W. The Application of Robust Regression to a Production Function Comparison – the Example of Swiss Corn. *The Open Agriculture Journal,* 2008, 2, pp. 90–98.
FOX, J. *Bootstrapping Robust regression models: Appendix to an R and S-PLUS. Companion to applied Regression* [online]. 2002. <http://cran.r-project.org/doc/contrib/FoxCompanion/appendix-bootstrapping.pdf>.
FREEDMAN, D. A. Bootstrapping regression models. *The Annals of Statistics,* 1981, 9(6), pp. 1218–1228.
GUELLEC, D., IOANNIDIS, E. Causes of Fluctuations in R&D Expenditures. A Quantitative Analysis. *OECD Economic Studies,* No. 29, 1997/II.
GUELLEC, D., POTTELSBERGHE DE LA POTTERIE, B. R&D and Productivity Growth: Panel Data Analysis of 16 OECD Countries. *OECD Economic Studies,* No. 33, 2001/II.
HAMPEL, F. R., RONCHETTI, E. M., ROUSSEEUW, P. J., STAHEL, W. A. *Robust Statistics. The Approach Based on Influence Functions.* New York: J. Willey, 1986.
HAMPEL, F. R. Some Aspects of Model Choice in Robust Statistics. In *Proceedings of the 44th Session of the ISI*, Book 2, 1983, pp. 767–771.
HUBERT, M., ROUSSEEUW, P. J., VAN AELST. High-Breakdown Robust Multivariate Methods. *Statistical Science*, 2008, 23(1), pp. 92–119.
KROLL. H., ZENKER, A., SCHUBERT, T. *An Analysis of the Development of R&D Expenditure at Regional Level in the Light of the 3% Target.* Luxembourg: Publications Office of the European Union, 2009.
OLIVE, D. J. Applications of robust distances for regression. *Technometrics.* 2002, 44(1), pp. 64–71.
RONCHETTI, E. Robust Model Selection in Regression. *Statistics & Probability Letters,* 1985, 3, pp. 21–23.
ROUSSEEUW, P. J. Least median of squares regression. *Journal of the American Statistical Association*, 1984, 79(388), pp. 871–880.
ROUSSEEUW, P. J., LEROY, A. M. *Robust Regression and Outlier Detection.* New Jersey: J. Willey, 2003.
ROUSSEEUW, P. J., VAN ZOMEREN, B. C. Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association,* 1990, 85(411), pp. 633–639.
RUPPERT, D., CARROLL, R. J. Trimmed Least Squares Estimation in the Linear Model. *Journal of the American Statistical Association,* 1990, 75, pp. 828–838.
SALIBIAN-BARRERA, M., ZAMAR, R. H. Bootstrapping robust estimates of regression. *The Annals of Statitics,* 2002, 30(2), pp. 556–582.
SAS 9.2 Help and documentation.
SOMMER, S., HUGGINS, R. M. Variable Selection Using the Wald Test and a Robust Cp. *Applied Statistics,* 1996, 45, pp. 15–29.
SPIŠÁKOVÁ, E. Influence of the Economic Crisis on the Fulfillment of the Target of Strategy Europe 2020 in the Area of Research and Development. In RAJMUND, M., eds. *Financial Aspects of Recent Trends in the Global Economy,* ASERS Publishing, 2013(1), No 1.
S-PLUS 6 Robust Library User's Guide. Seatle, Washington: Insightful Corporation, 2002.
STINE, R. An introduction to bootstrap methods, examples and ideas. *Sociological Methods and Research,* 1990, 18(2–3), pp. 243–291.
YOHAI, V. J. High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics,* 1987, 15(20), pp. 642–656.
ZAMAN, A., ROUSSEEUW, P. J., ORHAN, M. Econometric applications of high-breakdown robust regression techniques. *Economics Letters,* Vol. 71, 2000, pp. 1–8.
ZHANG, J. T. An empirical Analysis for National Energy R&D Expenditures. *International Journal of Global Energy Issues,* 2006, Vol. 26, No 1–2, pp. 141–159.

ANNEX: LIST OF INDICATORS

CPL Comparative Price Level (EU-27 =100%); (tec00120),

ER Employment rate total (the ratio of employed persons aged 20–64 and the total population of the same age group (t2020_10); (tsdec420),

GERD Gross domestic expenditure on R&D (total gross domestic expenditure on research and experimental development as a percentage of GDP; (t2020_20), (tsdec320),
GGD General government debt (percentage of GDP); (tsdde410),
HBA Households with broadband access to the Internet (percentage of all households); (tin00073),
HICP Harmonised indices of consumer prices (2005 = 100); (tec0027),
HRST Human Resources in Science and Technology (percentage of active population aged 25–64 years; (tsc00025),
HTE High-tech exports; (tin00140),
ILCS Individuals' level of computer skills (in percent) (tsdsc470),
IR Inflation rate (HICP); (tec00118),
IRUI Individuals regularly using the Internet (in percent; frequency of Internet access: once a week); (tin00091),
LLL Life-long learning (participation in education and training; percentage of people aged 25–64); (tsdsc440),
LPH Labour productivity per hour worked; (tec00117),
LPP Labour productivity per person employed; (tec00116),
LTU Long-term unemployment, total (annual average; percentage of active population); (tsdsc330),
PUSE Persons with upper secondary or tertiary education attainment (in percent), 25–64 years; (tps00065),
REER Real effective exchange rate (index, 2005 = 100); (tsdec330),
SRE Share of renewables in gross final energy consumption (tsdcc110);
UR Unemployment rate, total (percentage of the labour force); (tsdec450),
TEA Tertiary educational attainment, age group 30–34 (t2020_41),
TEAT Tertiary educational attainment, age group 25–64 (tps00065).

Alternative Means of Statistical Data Analysis: L-Moments and TL-Moments of Probability Distributions

Diana Bílková[1] | *University of Economics, Prague, Czech Republic*

Abstract

Moments and cumulants are commonly used to characterize the probability distribution or observed data set. The use of the moment method of parameter estimation is also common in the construction of an appropriate parametric distribution for a certain data set. The moment method does not always produce satisfactory results. It is difficult to determine exactly what information concerning the shape of the distribution is expressed by its moments of the third and higher order. In the case of small samples in particular, numerical values of sample moments can be very different from the corresponding values of theoretical moments of the relevant probability distribution from which the random sample comes. Parameter estimations of the probability distribution made by the moment method are often considerably less accurate than those obtained using other methods, particularly in the case of small samples. The present paper deals with an alternative approach to the construction of an appropriate parametric distribution for the considered data set using order statistics.

Keywords

Mikrocensus, L-moments and TL-moments of probability distribution, sample L-moments and TL-moments, probability density function, distribution function, quantile function, order statistics, income distribution

JEL code

C13, C46, C51, C52, C55, D31

INTRODUCTION

L-moments form the basis for a general theory which includes the summarization and description of theoretical probability distributions and obtained sample data sets, parameter estimation of theoretical probability distributions and hypothesis testing of parameter values for theoretical probability distributions. The theory of L-moments includes the established methods such as the use of order statistics and the Gini mean difference. It leads to some promising innovations in the area of measuring skewness and kurtosis of the distribution and provides relatively new methods of parameter estimation for an in-

[1] University of Economics, Prague, Faculty of Informatics and Statistics; Department of Statistics and Probability, W. Churchill Sq. 1938/4, 130 67 Prague 3, Czech Republic. E-mail: bilkova@vse.cz, phone: (+420)224095484.

dividual distribution. L-moments can be defined for any random variable whose expected value exists. The main advantage of L-moments over conventional moments is that they can be estimated by linear functions of sample values and are more resistant to the influence of sample variability. L-moments are more robust than conventional moments to the existence of outliers in the data, facilitating better conclusions made on the basis of small samples of the basic probability distribution. L-moments sometimes bring even more efficient parameter estimations of the parametric distribution than those estimated by the maximum likelihood method for small samples in particular, see Hosking (1990).

L-moments have certain theoretical advantages over conventional moments consisting in the ability to characterize a wider range of the distribution (i.e. range of values that the random variable can take including the extreme values). They are also more resistant and less prone to estimation bias, approximation by the asymptotic normal distribution being more accurate in finite samples, see Serfling (1980).

Let X be a random variable being distributed with the distribution function $F(x)$ and quantile function $x(F)$ and let $X_1, X_2, \ldots, X_n$ be a random sample of the sample size n from this distribution. Then $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ are order statistics of the random sample of the sample size n which comes from the distribution of the random variable X.

L-moments are analogous to conventional moments. They can be estimated on the basis of linear combinations of sample order statistics, i.e. L-statistics. L-moments are an alternative system describing the shape of the probability distribution.

1 METHODS AND METHODOLOGY

1.1 L-Moments of Probability Distributions

The issue of L-moments is discussed, for example, in Adamowski (2000) or Ulrych et al. (2000). Let X be a continuous random variable being distributed with the distribution function $F(x)$ and quantile function $x(F)$. Let $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ be order statistics of a random sample of the sample size n which comes from the distribution of the random variable X. L-moment of the r-th order of the random variable X is defined as:

$$\lambda_r = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot E(X_{r-j:r}), \quad r = 1, 2, \ldots . \tag{1}$$

An expected value of the r-th order statistic of the random sample of the sample size n has the form:

$$E(X_{r:n}) = \frac{n!}{(r-1)! \cdot (n-r)!} \cdot \int_0^1 x(F) \cdot [F(x)]^{r-1} \cdot [1 - F(x)]^{n-r} \mathrm{d}\, F(x). \tag{2}$$

If we substitute equation (2) into equation (1), after adjustments we obtain:

$$\lambda_r = \int_0^1 x(F) \cdot P^*_{r-1}[F(x)] \, \mathrm{d}\, F(x), \quad r = 1, 2, \ldots, \tag{3}$$

where:

$$P^*_r[F(x)] = \sum_{j=0}^{r} p^*_{r,j} \cdot [F(x)]^j \quad \text{a} \quad p^*_{r,j} = (-1)^{r-j} \cdot \binom{r}{j} \cdot \binom{r+j}{j}, \tag{4}$$

$P_r^*[F(x)]$ being the r-th shifted Legendre polynomial. Having substituted expression (2) into expression (1), we also obtained:

$$\lambda_r = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot \frac{r!}{(r-j-1)! \cdot j!} \cdot \int_0^1 x(F) \cdot [F(x)]^{r-j-1} \cdot [1 - F(x)]^j \, \mathrm{d}\, F(x), \quad r = 1, 2, \dots . \tag{5}$$

The letter "L" in "L-moments" indicates that the r-th L-moment λ_r is a linear function of the expected value of a certain linear combination of order statistics. The estimate of the r-th L-moment λ_r, based on the sample, is thus the linear combination of order data values, i.e. L-statistics. The first four L-moments of the probability distribution are now defined as:

$$\lambda_1 = E(X_{1:1}) = \int_0^1 x(F) \, \mathrm{d}\, F(x), \tag{6}$$

$$\lambda_2 = \frac{1}{2} E(X_{2:2} - X_{1:2}) = \int_0^1 x(F) \cdot [2\, F(x) - 1] \, \mathrm{d}\, F(x), \tag{7}$$

$$\lambda_3 = \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3}) = \int_0^1 x(F) \cdot \{6[F(x)]^2 - 6\, F(x) + 1\} \, \mathrm{d}\, F(x), \tag{8}$$

$$\lambda_4 = \frac{1}{4} E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) = \int_0^1 x(F) \cdot \{20[F(x)]^3 - 30[F(x)]^2 + 12[F(x)] - 1\} \, \mathrm{d}\, F(x). \tag{9}$$

The probability distribution can be specified by its L-moments even if some of its conventional moments do not exist, the opposite, however, is not true. It can be proved that the first L-moment λ_1 is a location characteristic, the second L-moment λ_2 being a variability characteristic. It is often desirable to standardize higher L-moments λ_r, $r \geq 3$, so that they can be independent of specific units of the random variable X. The ratio of L-moments of the r-th order of the random variable X is defined as:

$$\tau_r = \frac{\lambda_r}{\lambda_2}, \quad r = 3, 4, \dots . \tag{10}$$

We can also define the function of L-moments which is analogous to the classical coefficient of variation, i.e. the so called L-coefficient of variation:

$$\tau = \frac{\lambda_2}{\lambda_1}. \tag{11}$$

The ratio of L-moments τ_3 is a skewness characteristic, the ratio of L-moments τ_4 being a kurtosis characteristic of the corresponding probability distribution. Main properties of the probability distribution are very well summarized by the following four characteristics: L-location λ_1, L-variability λ_2, L-skewness τ_3 and L-kurtosis τ_4. L-moments λ_1 and λ_2, the L-coefficient of variation τ and ratios of L-moments τ_3 and τ_4 are the most useful characteristics for the summarization of the probability distribution. Their main properties are existence (if the expected value of the distribution is finite, then all its L-moments exist) and uniqueness (if the expected value of the distribution is finite, then L-moments define the only distribution, i.e. no two distinct distributions have the same L-moments).

Using equations (6)–(9) and (10), we obtain both the expressions for L-moments and L-moments ratios for lognormal and generalized Pareto probability distributions, see Table 1.

Table 1 Formulas for distribution or quantile functions, L-moments and their ratios for lognormal and generalized Pareto probability distributions

Distribution	Distribution function *F*(*x*) or quantile function *x*(*F*)	L-moments and ratios of L-moments
Logo normal	$F(x) = \Phi\left\{\frac{\ln[x(F)-\xi]-\mu}{\sigma}\right\}$	$\lambda_1 = \xi + \exp\left(\mu + \frac{\sigma^2}{2}\right)$ $\lambda_2 = \exp\left(\mu + \frac{\sigma^2}{2}\right)\cdot \operatorname{erf}\left(\frac{\sigma}{2}\right)$ $\tau_3 = 6\pi^{-\frac{1}{2}}\cdot\frac{\int_0^{\frac{\sigma}{2}} \operatorname{erf}\left(\frac{x}{\sqrt{3}}\right)\cdot\exp(-x^2)\,\mathrm{d}x}{\operatorname{erf}\left(\frac{\sigma}{2}\right)}$
Generalized Pareto	$x(F) = \xi + \alpha\cdot\frac{1-[1-F(x)]^k}{k}$	$\lambda_1 = \xi + \frac{\alpha}{1+k}$ $\lambda_2 = \frac{\alpha}{(1+k)\cdot(2+k)}$ $\tau_3 = \frac{1-k}{3+k}$ $\tau_4 = \frac{(1-k)\cdot(2-k)}{(3+k)\cdot(4+k)}$

Source: Hosking (1990); own research

1.2 Sample L-Moments

L-moments are usually estimated by a random sample obtained from an unknown distribution. Since the r-th L-moment λ_r is the function of the expected values of order statistics of a random sample of the sample size r, it is natural to estimate it using the so-called U-statistic, i.e. the corresponding function of sample order statistics (averaged over all subsets of the sample size r, which may be formed from the obtained random sample of the sample size n).

Let $x_1, x_2, \ldots, x_n$ be the sample and $x_{1:n} \le x_{2:n} \le \ldots \le x_{n:n}$ the ordered sample. Then the r-th sample L-moment can be written as:

$$l_r = \binom{n}{r}^{-1} \sum_{1\le i1 < i2 < \ldots < ir \le n} \ldots \sum \frac{1}{r}\cdot\sum_{j=0}^{r-1}(-1)^j\cdot\binom{r-1}{j}\cdot x_{i r-j:n}, \quad r = 1, 2, \ldots, n. \tag{12}$$

Hence the first four sample L-moments have the form:

$$l_1 = \frac{1}{n}\cdot\sum_i x_i, \tag{13}$$

$$l_2 = \frac{1}{2}\cdot\binom{n}{2}^{-1}\cdot\sum_{i>j}\sum (x_{i:n} - x_{j:n}), \tag{14}$$

$$l_3 = \frac{1}{3} \cdot \binom{n}{3}^{-1} \cdot \sum_{i>j>k}\sum\sum (x_{i:n} - 2x_{j:n} + x_{k:n}), \tag{15}$$

$$l_4 = \frac{1}{4} \cdot \binom{n}{4}^{-1} \cdot \sum_{i>j>k>l}\sum\sum\sum (x_{i:n} - 3x_{j:n} + 3x_{k:n} - x_{l:n}). \tag{16}$$

U-statistics are widely used especially in nonparametric statistics. Their positive properties are the absence of bias, asymptotic normality and a slight resistance due to the influence of outliers, see Hosking (1990).

When calculating the r-th sample L-moment, it is not necessary to repeat the process over all sub-sets of the sample size r, since this statistic can be expressed directly as a linear combination of order statistics of a random sample of the sample size n.

If we assume an estimate of $E(X_{r:r})$ obtained with the use of U-statistics, it can be written as $r \cdot b_{r-1}$, where:

$$b_r = \frac{1}{n} \cdot \binom{n-1}{r}^{-1} \cdot \sum_{j=r+1}^{n} \binom{j-1}{r} \cdot x_{j:n}, \tag{17}$$

namely:

$$b_0 = \frac{1}{n} \cdot \sum_{j=1}^{n} x_{j:n}, \tag{18}$$

$$b_1 = \frac{1}{n} \cdot \sum_{j=2}^{n} \frac{(j-1)}{(n-1)} \cdot x_{j:n}, \tag{19}$$

$$b_2 = \frac{1}{n} \cdot \sum_{j=3}^{n} \frac{(j-1)\cdot(j-2)}{(n-1)\cdot(n-2)} \cdot x_{j:n}, \tag{20}$$

and so generally:

$$b_r = \frac{1}{n} \cdot \sum_{j=r+1}^{n} \frac{(j-1)\cdot(j-2)\cdot\ldots\cdot(j-r)}{(n-1)\cdot(n-2)\cdot\ldots\cdot(n-r)} \cdot x_{j:n}. \tag{21}$$

Thus the first sample L-moments can be written as:

$$l_1 = b_0, \tag{22}$$

$$l_2 = 2b_1 - b_0, \tag{23}$$

$$l_3 = 6b_2 - 6b_1 + b_0, \tag{24}$$

$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0, \tag{25}$$

We can therefore write generally:

$$l_{r+1} = \sum_{k=0}^{r} p^{*}_{r,k} \cdot b_k, \quad r = 0, 1, \ldots, n-1, \tag{26}$$

where:

$$p^{*}_{r,k} = (-1)^{r-k} \cdot \binom{r}{k} \cdot \binom{r+k}{k} = \frac{(-1)^{r-k} \cdot (r+k)!}{(k!)^2 \cdot (r-k)!}. \tag{27}$$

Sample L-moments are used in a similar way as sample conventional L-moments, summarizing the basic properties of the sample distribution, which are the location (level), variability, skewness and kurtosis. Thus, sample L-moments allow an estimation the corresponding properties of the probability distribution from which the sample originates and can be used in estimating the parameters of the relevant probability distribution. We often prefer L-moments to conventional moments within such applications, since sample L-moments – as the linear functions of sample values – are less sensitive to sample variability or measurement errors in extreme observations than conventional moments. L-moments therefore lead to more accurate and robust estimates of characteristics or parameters of the basic probability distribution.

Sample L-moments have been used previously in statistics, but not as part of a unified theory. The first sample L-moment l_1 is a sample L-location (sample average), the second sample L-moment l_2 being a sample L-variability. The natural estimation of L-moments (10) ratio is the sample ratio of L-moments:

$$t_r = \frac{l_r}{l_2}, \quad r = 3, 4, \ldots. \tag{28}$$

Hence t_3 is a sample L-skewness and t_4 is a sample L-kurtosis. Sample ratios of L-moments t_3 and t_4 may be used as the characteristics of skewness and kurtosis of a sample data set.

Table 2 Formulas for parameter estimations made by the method of L-moments of lognormal and generalized Pareto probability distributions

Distribution	Parameter estimation
Logo normal	$z = \sqrt{\frac{8}{3}} \cdot \Phi^{-1}\left(\frac{1+t_3}{2}\right)^{2)}$ $\hat{\sigma} = 0.999\,281 z - 0.006\,118\, z^3 + 0.000\,127\, z^5$ $\hat{\mu} = \ln \frac{l_2}{\mathrm{erf}\left(\frac{\sigma}{2}\right)} - \frac{\hat{\sigma}^2}{2}$ $\hat{\xi} = l_1 - \exp\left(\hat{\mu} + \frac{\hat{\sigma}^2}{2}\right)$
Generalized Pareto	(ξ známé) $\hat{k} = \frac{l_1}{l_2} - 2$ $\hat{\alpha} = (1+\hat{k}) \cdot l_1$

Source: Hosking (1990); own research

The Gini mean difference relates both to sample L-moments, having the form of:

$$G = \binom{n}{2}^{-1} \cdot \sum_{i>j} \sum (x_{i:n} - x_{j:n}), \tag{29}$$

and the Gini coefficient which depends only on a single parameter σ in the case of the two-parametric lognormal distribution, depending, however, on the values of all three parameters in the case of the three-parametric lognormal distribution. Table 2 presents the expressions for parameter estimations of lognormal and generalized Pareto probability distributions obtained using the method of L-moments. For more details see, for example, Bílková (2010), Bílková (2011), Bílková (2012), Bílková, Malá (2012), Hosking (1990) or Kyselý, Picek (2007).

1.3 TL-Moments of Probability Distributions

An alternative robust version of L-moments is introduced in this subchapter. The modification is called "trimmed L-moments" and it is termed TL-moments. The expected values of order statistics of a random sample in the definition of L-moments of probability distributions are replaced with those of a larger random sample, its size growing correspondingly to the extent of the modification, as shown below.

Certain advantages of TL-moments outweigh those of conventional L-moments and central moments. TL-moment of the probability distribution may exist despite the non-existence of the corresponding L-moment or central moment of this probability distribution, as it is the case of the Cauchy distribution. Sample TL-moments are more resistant to outliers in the data. The method of TL-moments is not intended to replace the existing robust methods but rather supplement them, particularly in situations when we have outliers in the data.

In this alternative robust modification of L-moments, the expected value $E(X_{r-j:r})$ is replaced with the expected value $E(Xr + t_1 - j{:}r + t_1 + t_2)$. Thus, for each r, we increase the sample size of a random sample from the original r to $r + t_1 + t_2$, working only with the expected values of these r modified order statistics $Xt_1 + 1{:}r + t_1 + t_2$, $Xt_1 + 2{:}r + t_1 + t_2$, …, $Xt_1 + r{:}r + t_1 + t_2$ by trimming the smallest t_1 and largest t_2 from the conceptual random sample. This modification is called the r-th trimmed L-moment (TL-moment) and marked as $\lambda_r^{(t_1,t_2)}$. Thus, TL-moment of the r-th order of the random variable X is defined as:

$$\lambda_r^{(t_1,t_2)} = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot E(X_{r+t_1-j:r+t_1+t_2}), \quad r = 1, 2, \dots . \tag{30}$$

It is evident from the expressions (30) and (1) that TL-moments are reduced to L-moments, where $t_1 = t_2 = 0$. Although we can also consider applications where the adjustment values are not equal, i.e. $t_1 \neq t_2$, we will focus here only on the symmetric case $t_1 = t_2 = t$. Then the expression (30) can be rewritten:

$$\lambda_r^{(t)} = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot E(X_{r+t-j:r+2t}), \quad r = 1, 2, \dots . \tag{31}$$

Thus, for example, $\lambda_1^{(t)} = E(X_{1+t:1+2t})$ is the expected value of the median of the conceptual random sample of $1 + 2t$ size. It is necessary to note that $\lambda_1^{(t)}$ is equal to zero for distributions that are symmetrical around zero.

For $t = 1$, the first four TL-moments have the form:

$$\lambda_1^{(1)} = E(X_{2:3}), \tag{32}$$

$$\lambda_2^{(1)} = \frac{1}{2} E(X_{3:4} - X_{2:4}), \tag{33}$$

$$\lambda_3^{(1)} = \frac{1}{3} E(X_{4:5} - 2X_{3:5} + X_{2:5}), \tag{34}$$

$$\lambda_4^{(1)} = \frac{1}{4} E(X_{5:6} - 3X_{4:6} + 3X_{3\ 6} - X_{2:6}). \tag{35}$$

The measurements of location, variability, skewness and kurtosis of the probability distribution analogous to conventional L-moments (6)–(9) are based on $\lambda_1^{(1)}, \lambda_2^{(1)}, \lambda_3^{(1)}$ a $\lambda_4^{(1)}$.

The expected value $E(X_{r:n})$ can be written using the formula (2). With the use of the equation (2), we can express the right side of the equation (31) again as:

$$\lambda_r^{(t)} = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot \frac{(r+2t)!}{(r+t-j-1)! \cdot (t+j)!} \cdot \int_0^1 x(F) \cdot [F(x)]^{r+t-j-1} \cdot [1-F(x)]^{t+j} \, \mathrm{d}F(x), r = 1, 2, \ldots. \tag{36}$$

It is necessary to point out that $\lambda_r^{(0)} = \lambda_r$ represents a normal r-th L-moment with no respective adjustments.

Expressions (32)–(35) for the first four TL-moments ($t = 1$) may be written in an alternative way as:

$$\lambda_1^{(1)} = 6 \cdot \int_0^1 x(F) \cdot [F(x)] \cdot [1-F(x)] \, \mathrm{d}F(x), \tag{37}$$

$$\lambda_2^{(1)} = 6 \cdot \int_0^1 x(F) \cdot [F(x)] \cdot [1-F(x)] \cdot [2F(x)-1] \, \mathrm{d}F(x), \tag{38}$$

$$\lambda_3^{(1)} = \frac{20}{3} \cdot \int_0^1 x(F) \cdot [F(x)] \cdot [1-F(x)] \cdot \{5[F(x)]^2 - 5F(x) + 1\} \, \mathrm{d}F(x), \tag{39}$$

$$\lambda_4^{(1)} = \frac{15}{2} \cdot \int_0^1 x(F) \cdot [F(x)] \cdot [1-F(x)] \cdot \{14[F(x)]^3 - 21[F(x)]^2 + 9[F(x)] - 1\} \, \mathrm{d}F(x). \tag{40}$$

The distribution can be determined by its TL-moments, even though some of its L-moments or conventional moments do not exist. For example, $\lambda_1^{(1)}$ (the expected value of the median of a conceptual random sample of sample size three) exists for the Cauchy distribution, despite the non-existence of the first L-moment λ_1.

TL-skewness $\tau_3^{(t)}$ and TL-kurtosis $\tau_4^{(t)}$ can be defined analogously as L-skewness τ_3 and L-kurtosis τ_4

$$\tau_3^{(t)} = \frac{\lambda_3^{(t)}}{\lambda_2^{(t)}}, \tag{41}$$

$$\tau_4^{(t)} = \frac{\lambda_4^{(t)}}{\lambda_2^{(t)}}, \tag{42}$$

1.4 Sample TL-Moments

Let $x_1, x_2, \ldots, x_n$ be a sample and $x_{1:n} \le x_{2:n} \le \ldots \le x_{n:n}$ an order sample. The expression:

$$\hat{E}(X_{j+1:j+l+1}) = \frac{1}{\binom{n}{j+l+1}} \cdot \sum_{i=1}^{n} \binom{i-1}{j} \cdot \binom{n-i}{l} \cdot x_{i:n} \tag{43}$$

is considered to be an unbiased estimate of the expected value of the $(j+1)$-th order statistic $X_{j+1:j+l+1}$ in the conceptual random sample of sample size $(j+l+1)$. Now we will assume that in the definition of TL-moment $\lambda_r^{(t)}$ in (31), the expression $E(X_{r+t-j:r+2t})$ is replaced by its unbiased estimate:

$$\hat{E}(X_{r+t-j:r+2t}) = \frac{1}{\binom{n}{r+2t}} \cdot \sum_{i=1}^{n} \binom{i-1}{r+t-j-1} \cdot \binom{n-i}{t+j} \cdot x_{i:n}, \tag{44}$$

which is obtained by assigning $j \to r+t-j-1$ a $l \to t+j$ in (43). Now we get the <u>r-th sample TL-moment:</u>

$$l_r^{(t)} = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot \hat{E}(X_{r+t-j:r+2t}), \quad r = 1, 2, \ldots, n-2t, \tag{45}$$

i.e.:

$$l_r^{(t)} = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot \frac{1}{\binom{n}{r+2t}} \cdot \sum_{i=1}^{n} \binom{i-1}{r+t-j-1} \cdot \binom{n-i}{t+j} \cdot x_{i:n}, \quad r = 1, 2, \ldots, n-2t, \tag{46}$$

which is an unbiased estimate of the r-th TL-moment $\lambda_r^{(t)}$. Let us note that for each $j = 0, 1, \ldots, r-1$, the values $x_{i:n}$ in (46) are not equal to zero only for $r+t-j \le i \le n-t-j$, taking combination numbers into account. A simple adjustment of equation (46) provides an alternative linear form:

$$l_r^{(t)} = \frac{1}{r} \cdot \sum_{i=r+t}^{n-t} \left[\frac{\sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \binom{i-1}{r+t-j-1} \cdot \binom{n-i}{t+j}}{\binom{n}{r+2t}} \right] \cdot x_{i:n}. \tag{47}$$

For $r = 1$, for example, we obtain for the first sample TL-moment:

$$l_1^{(t)} = \sum_{i=t+1}^{n-t} w_{i:n}^{(t)} \cdot x_{i:n}, \tag{48}$$

where the weights are given by:

$$w_{i:n}^{(t)} = \frac{\binom{i-1}{t} \cdot \binom{n-i}{t}}{\binom{n}{2t+1}}. \tag{49}$$

The above results can be used for the estimation of TL-skewness $\tau_4^{(t)}$ and TL-kurtosis $\tau_4^{(t)}$ by simple ratios:

$$t_3^{(t)} = \frac{l_3^{(t)}}{l_2^{(t)}}, \tag{50}$$

$$t_4^{(t)} = \frac{l_4^{(t)}}{l_2^{(t)}}. \tag{51}$$

We can choose $t = n\alpha$, representing the size of the adjustment from each end of the sample, where α is a certain ratio, where $0 \le \alpha < 0{,}5$.

Table 3 contains the expressions for TL-moments and their ratios as well as those for parameter estimations of logistic and Cauchy probability distributions obtained employing the method of TL-moments ($t = 1$); for more, see, e.g. Elamir, Seheult (2003).

Table 3 Formulas for TL-moments and their ratios and parameter estimations made by the method of TL-moments of logistic and Cauchy probability distributions ($t = 1$)

Distribution	TL-moments and ratios of TL-moments	Parameter estimation
Logistic	$\lambda_1^{(1)} = \mu$ $\lambda_2^{(1)} = 0{,}500\sigma$ $\tau_3^{(1)} = 0$ $\tau_4^{(1)} = 0{,}083$	$\hat{\mu} = l_1^{(1)}$ $\hat{\sigma} = 2l_2^{(1)}$
Cauchy	$\lambda_1^{(1)} = \mu$ $\lambda_2^{(1)} = 0.698\sigma$ $\tau_3^{(1)} = 0$ $\tau_4^{(1)} = 0{,}343$	$\hat{\mu} = l_1^{(1)}$ $\hat{\sigma} = \frac{l_2^{(1)}}{0.698}$

Source: Elamir, Seheult (2003); own research

1.5 Maximum Likelihood Method

Let a random sample of sample size n come from the three-parametric lognormal distribution with a probability density function:

$$\begin{aligned} f(x; \mu, \sigma^2, \theta) &= \frac{1}{\sigma \cdot (x-\theta) \cdot \sqrt{2\pi}} \cdot \exp\left[-\frac{[\ln(x-\theta)-\mu]^2}{2\sigma^2}\right], \quad x > \theta, \\ &= 0, \end{aligned} \tag{52}$$

where $-\infty < \mu < \infty$, $\sigma^2 > 0$, $-\infty < \theta < \infty$ are parameters. The three-parametric lognormal distribution is described in detail, for example, in Bílková (2010), Bílková (2011) and Bílková (2012).

The likelihood function then has the form:

$$L(\boldsymbol{x};\mu,\sigma^2,\theta)=\prod_{i=1}^{n} f(x_i;\mu,\sigma^2,\theta)=$$
$$=\frac{1}{(\sigma^2)^{n/2}\cdot(2\pi)^{n/2}\cdot\prod\limits_{i=1}^{n}(x_i-\theta)}\cdot\exp\left\{\sum_{i=1}^{n}-\frac{[\ln(x_i-\theta)-\mu]^2}{2\sigma^2}\right\}. \tag{53}$$

We determine the natural logarithm of the likelihood function:

$$\ln L(\boldsymbol{x};\mu,\sigma^2,\theta)=\sum_{i=1}^{n}-\frac{[\ln(x_i-\theta)-\mu]^2}{2\sigma^2}-\frac{n}{2}\cdot\ln\sigma^2-\frac{n}{2}\cdot\ln(2\pi)-\sum_{i=1}^{n}\ln(x_i-\theta). \tag{54}$$

We make the first partial derivatives of the likelihood function logarithm according to μ and σ^2 equal to zero, obtaining a system of likelihood equations:

$$\frac{\partial\ln L(\boldsymbol{x};\mu,\sigma^2,\theta)}{\partial\mu}=\frac{\sum\limits_{i=1}^{n}[\ln(x_i-\theta)-\mu]}{\sigma^2}=0, \tag{55}$$

$$\frac{\partial\ln L(\boldsymbol{x};\mu,\sigma^2,\theta)}{\partial\sigma^2}=\frac{\sum\limits_{i=1}^{n}[\ln(x_i-\theta)-\mu]^2}{2\sigma^4}-\frac{n}{2\sigma^2}=0. \tag{56}$$

After adjustment we obtain maximum likelihood estimations of parameters μ and σ^2 for the parameter θ:

$$\hat{\mu}(\theta)=\frac{\sum\limits_{i=1}^{n}\ln(x_i-\theta)}{n}, \tag{57}$$

$$\hat{\sigma}^2(\theta)=\frac{\sum\limits_{i=1}^{n}[\ln(x_i-\theta)-\hat{\mu}(\theta)]^2}{n}. \tag{58}$$

If the value of the parameter θ is known, we get maximum likelihood estimates of the remaining two parameters of the three-parametric lognormal distribution using equations (57) and (58). However, if the value of the parameter θ is unknown, the problem is more complicated. It has been proved that if the parameter θ gets closer to $\min\{X_1, X_2, \ldots, X_n\}$, then the likelihood function approaches infinity. The maximum likelihood method is also often combined with the Cohen method, where the smallest sample value is made equal to $100\times(n+1)^{-1}$% quantile:

$$x_{\min}^{V}=\hat{\theta}+\exp(\hat{\mu}+\hat{\sigma}\cdot u_{(n+1)^{-1}}). \tag{59}$$

Equation (59) is then combined with the system of equations (57) and (58).

For the solution of maximum likelihood equations (57) and (58), it is also possible to use $\hat{\theta}$ satisfying the equation:

$$\sum_{i=1}^{n}(x_i-\hat{\theta}) + \frac{\sum_{i=1}^{n}\frac{z_i^{\backslash}}{(x_i-\hat{\theta})}}{\hat{\sigma}(\hat{\theta})} = 0, \tag{60}$$

where:

$$z_i^{\backslash} = \frac{\ln(x_i-\hat{\theta}) - \hat{\mu}(\hat{\theta})}{\hat{\sigma}(\hat{\theta})}, \tag{61}$$

where $\hat{\mu}(\hat{\theta})$ and $\hat{\sigma}(\hat{\theta})$ comply with equations (57) and (58), the parameter θ being replaced by $\hat{\theta}$. We may also obtain the bounds of variances:

$$n \cdot D(\hat{\theta}) = \frac{\sigma^2 \cdot \exp(2\mu)}{\omega \cdot \left[\omega \cdot (1+\sigma^2) - 2\sigma^2 - 1\right]}, \tag{62}$$

$$n \cdot D(\hat{\mu}) = \frac{\sigma^2 \cdot [\omega \cdot (1+\sigma^2) - 2\sigma^2]}{\omega \cdot (1+\sigma^2) - 2\sigma^2 - 1}, \tag{63}$$

$$n \cdot D(\hat{\sigma}) = \frac{\sigma^2 \cdot [\omega \cdot (1+\sigma^2) - 1]}{\omega \cdot (1+\sigma^2) - 2\sigma^2 - 1}. \tag{64}$$

2 RESULTS

L-moments method used to be employed in hydrology, climatology and meteorology in the research of extreme precipitation, see, e.g. Kyselý, Picek (2007), having mostly used smaller data sets. This study presents applications of L-moments and TL-moments to large sets of economic data, Table 4 showing the sample sizes of obtained household sample sets. Researched sampled sets of households constitute a reprezentative sample of the study population. The research variable is the net annual household income per capita (in CZK) in the Czech Republic (nominal income). The data collected by the Czech Statistical Office come from the Microcensus survey spanning the years 1992, 1996 and 2002. In total, 72 income distributions were analyzed – for all households in the Czech Republic as well as with the use of particular criteria: gender, region (Bohemia and Moravia), social group, municipality size, age and the highest educational attainment. The households are divided into subsets according to their heads – mostly men. The head of household is always a man in two-parent families (a husband-and-wife or cohabitee type), regardless of the economic activity. In lone-parent families (a one-parent-with-children type) and non-family households whose members are related neither by marriage (partnership) nor parent-child relationship, a crucial criterion for determining the head of household is the economic activity, another aspect being the amount of money income of individual household members. The former criterion also applies in the case of more complex household types, for instance, in joint households of more two-parent families.

Three-parametric lognormal distribution is here used as a basic theoretical probability distribution. Experience shows that the use of three-parametric lognormal curve as a model of income distribution is sufficient for global income models on a national scale and for income models arised using very gross classification with large sample sizes, see Hátle et al. (1975).

Table 4 Sample sizes of income distributions

	1992	1996	2002
Sample size	16 233	28 148	7 973

Source: Own research

Parameters of three-parametric lognormal curves were estimated simultaneously, three methods of parametric estimation having been employed – namely those of TL-moments, L-moments and maximum likelihood, their accuracy being compared to each other with the use of a common test criterion:

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - n\pi_i)^2}{n\pi_i}, \tag{65}$$

where n_i are the observed frequencies in particular income intervals, π_i are theoretical probabilities of a statistical unit belonging to the i-th interval, n is the total sample size of a corresponding statistical set, $n \times \pi_i$ are theoretical frequencies in particular income intervals, $i = 1, 2, ..., k$, and k is the number of intervals.

However, the appropriateness of a model curve for the income distribution is not a common mathematical and statistical issue encompassing tests of the null hypothesis.

H_0: The samplecomes from the assumed theoretical distribution

against the alternative hypothesis

H_1: non H_0,

since large sample sizes occur frequently in goodness of fit tests in the case of the income distribution, and hence the tests would mostly lead to the rejection of the null hypothesis. This results not only from a high power of the test at a chosen significance level, enabling it to indicate the slightest divergences between the actual income distribution and the model, but also from the test construction itself.

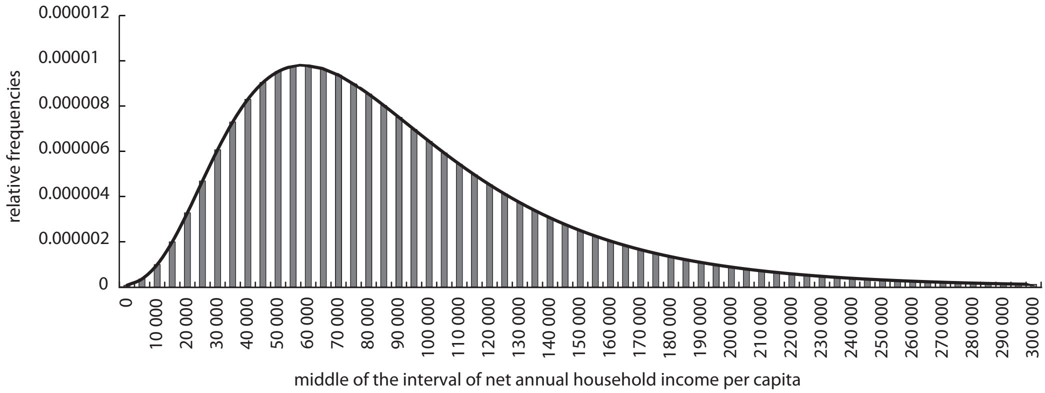
Not focusing, in fact, on small divergences, we are satisfied with a rough agreement of the model with the reality, the model (curve) being simply "borrowed". In this respect, only tentative conclusions can be drawn from the use of the test criterion χ^2. We have to assess the suitability of the model subjectively to some extent, relying on experience and logical analysis.

The value of $\alpha = 0.25$ from the middle of the interval $0 \leq \alpha < 0,5$ was used in this research. With only minor exceptions, the TL-moments method produced the most accurate results. L-moments was the second most effective method in more than half of the cases, the differences between this method and that of maximum likelihood not being significant enough as far as the number of cases, when the former gave better results than the latter. Table 5 represents distinctive outcomes for all 72 income distributions, showing the results for the total household sets in the Czech Republic. Apart from the estimated parameter values of the three-parametric lognormal distribution, which were obtained having simultaneously employed TL-moments, L-moments and maximum likelihood methods, Table 5 contains the values of the test criterion (65), indicating that the L-moments method produced – in two out of three cases – more accurate results than the maximum likelihood method, the most accurate outcomes in all three cases being produced by the TL-moments method.

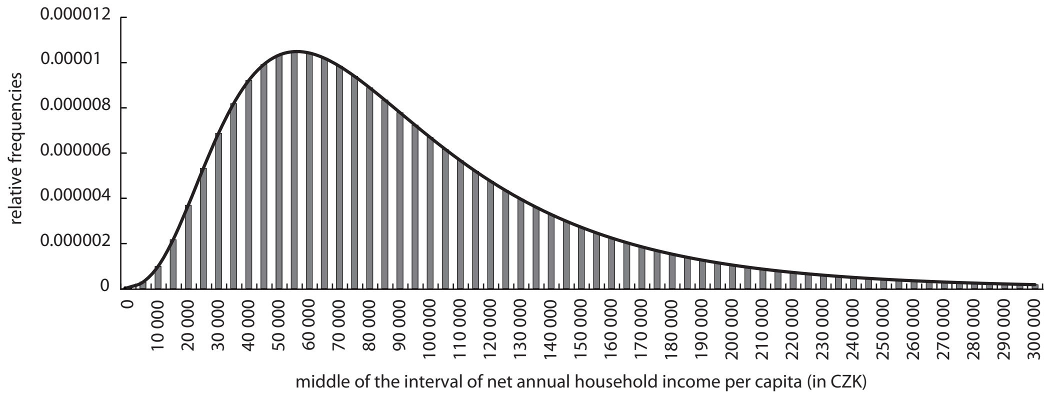
For the year 1992, an estimate of the value of the parameter θ (the beginning of the distribution, theoretical minimum) made by the maximum likelihood method is negative. This, however, may not interfere with good agreement between the model and the real distribution since the curve has initially a close contact with the horizontal axis.

Figure 1 Histograms of employees by net annual household income per capita with parameters of three-parametric lognormal curves estimated by the method of TL-moments method of L-moments and maximum likelihood method in 1992

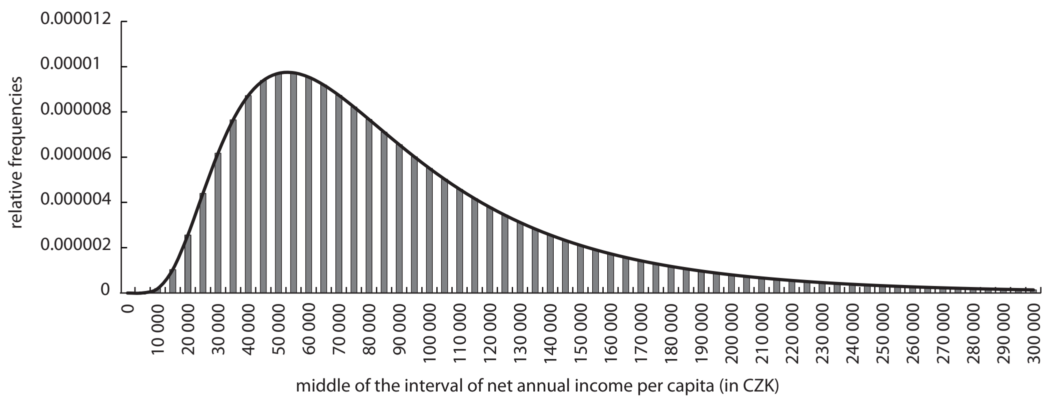
a. Method of TL-moments



b. Method of L-moments

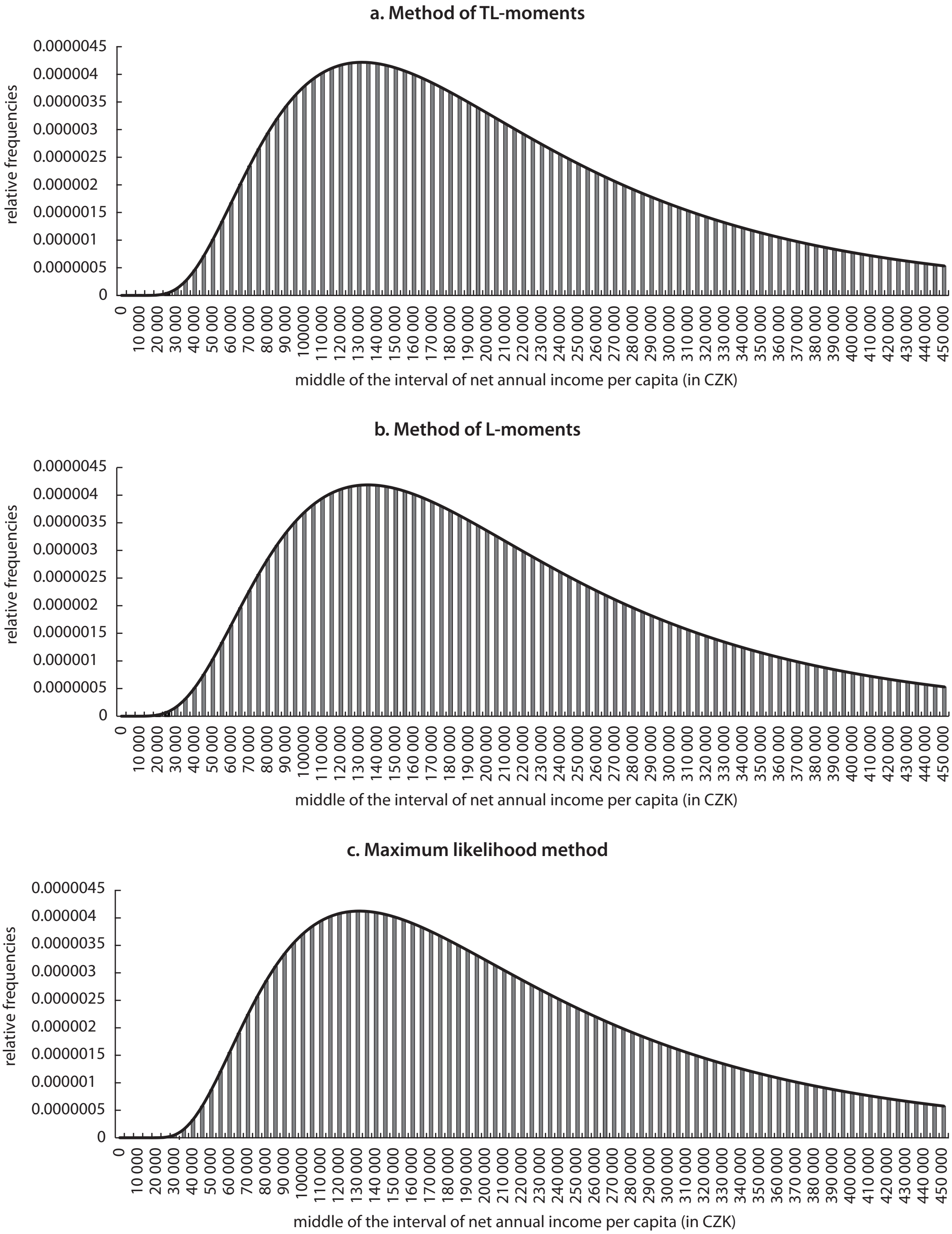


c. Maximum likelihood method



Source: Own research

Figure 2 Histograms of employees by net annual household income per capita with parameters of three-parametric lognormal curves estimated by the method of TL-moments method of L-moments and maximum likelihood method in 2002



Source: Own research

Table 5 Parameter estimations of three-parametric lognormal curves obtained using three various methods of point parameter estimation and the value of χ2 criterion

Year	Method of TL-moments			Method of L-moments			Maximum likelihood method		
	μ	σ2	θ	μ	σ2	θ	μ	σ2	θ
1992	9.722	0.521	14 881	9.696	0.700	14 491	10.384	0.390	–325
1996	10.334	0.573	25 981	10.343	0.545	25 362	10.995	0.424	52 231
2002	10.818	0.675	40 183	10.819	0.773	37 685	11.438	0.459	73 545

Year	Criterion χ2	Criterion χ2	Criterion χ2
1992	739.512	811.007	1 227.325
1996	1 503.878	1 742.631	2 197.251
2002	998.325	1 535.557	1 060.891

Source: Own research

Figures 1–2 allow us to compare the methods in terms of histogram of employees by net annual household income per capita with parameters of three-parametric lognormal curves estimated using various methods of parameter estimation in the given years (1992 and 2002) for the whole set of all households in the Czech Republic. It is clear from these figures that the methods of TL-moments and L-moments produce very similar results, while the histogram with the parameters estimated by the maximum likelihood method differs greatly from the histograms constructed using TL-moments and L-moments methods respectively.

A comparison of the accuracy of the three methods of point parameter estimation is also provided by Table 6. It shows the development of the sample median and theoretical medians of the lognormal distribution with the parameters estimated using the methods of TL-moments, L-moments and maximum likelihood for the whole set of households in the Czech Republic over the research period. This table also shows the differences between the theoretical and corresponding sample medians. It is also obvious from this table that the difference between the theoretical and sample medians is the smallest for the method of TL-moments, the method of L-moments follows and the maximum likelihood method is the least accurate.

Table 6 Theoretical medians obtained using the various method of parametric estimation, sample medians and the difference between the theoretical and sample median

Year	Median				Difference		
	Method of TL-moments	Method of L-moments	Maximum likelihood method	Sample median	Method of TL-moments	Method of L-moments	Maximum likelihood method
1992	30 743	31 562	32 013	31 000	–257	562	1 013
1996	56 742	56 401	59 628	57 700	–958	–1 299	1 928
2002	90 094	87 646	92 855	89 204	890	–1 558	3 651

Source: Own research

CONCLUSION

A relatively new class of moment characteristics of probability distributions has been introduced in the present paper. They are the characteristics of the location (level), variability, skewness and kurtosis of probability distributions constructed with the use of L-moments and TL-moments that represent a robust extension of L-moments. The very L-moments were implemented as a more robust alternative to classical moments of probability distributions. L-moments and their estimates, however, are lacking in some robust features that are associated with TL-moments.

Sample TL-moments are the linear combinations of sample order statistics assigning zero weight to a predetermined number of sample outliers. They are unbiased estimates of the corresponding TL-moments of probability distributions. Some theoretical and practical aspects of TL-moments are still the subject of both current and future research. The efficiency of TL-statistics depends on the choice of α, for example, $l_1^{(0)}, l_1^{(1)}, l_1^{(2)}$ have the smallest variance (the highest efficiency) among other estimates for random samples from the normal, logistic and double exponential distribution.

The above methods as well as other approaches, e.g. Marek (2011) or Marek, Vrabec (2013), can be also adapted for modelling the wage distribution and other economic data analysis.

ACKNOWLEDGEMENT

This paper was subsidized by the funds of institutional support of a long-term conceptual advancement of science and research number IP400040 at the Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic.

References

ADAMOWSKI, K. Regional Analysis of Annual Maximum and Partial Duration Flood Data by Nonparametric and L-moment Methods. *J. Hydrol.*, 2000, Vol. 229, pp. 219–231. ISSN 0022-1694.

BÍLKOVÁ, D. Estimating Parameters of Lognormal Distribution Using the Method of L-Moments. *Research Journal of Economics, Business and ICT,* 2011, Vol. 4, No. 1, pp. 4–9. ISSN 2045-3345.

BÍLKOVÁ, D. Modelling of Wage and Income Distributions Using the Method of L-Moments. *Journal of Mathematics and System Science*, 2012, Vol. 2, No. 1, pp. 13–19. ISSN 2159-5291.

BÍLKOVÁ, D. Modelling of Wage Distributions Using the Method of L-Moments. Paper presented at *AMSE – Applications of Mathematics and Statistics in Economy* held on 25–28 August 2010, Demänovská Dolina, pp. 16–30. ISBN 978-80-89438-02-0.

BÍLKOVÁ, D. L-Moments and Their Use in Modeling the Distribution of Income and Wage. Paper presented at *ISI* held on 21–26 August 2011, Dublin, flashdisk, pp. 1–6.

BÍLKOVÁ, D. Lognormal Distribution and Using L-Moment Method for Estimating Its Parameters [online]. *International Journal of Mathematical Models and Methods in Applied Sciences,* 2012, Vol. 6, No. 1, pp. 30–44. ISSN 1998-0140. <http://www.naun.org/journals/m3as/17-079.pdf>.

BÍLKOVÁ, D. Lognormal Distribution Parameter Estimating Using L-Moments. *Journal of Mathematics and Technology,* 2012, Vol. 3, No. 1, pp. 33–51. ISSN: 2078-0257.

BÍLKOVÁ, D. Modeling of Income and Wage Distribution Using the Method of L-Moments of Parameter Estimation. Paper presented at *International Days of Statistics and Economics* at VŠE held on 22–23 September 2011, Prague, pp. 1–10. ISBN 978-80-86175-72-0.

BÍLKOVÁ, D. Three-Parametric Lognormal Distribution and Estimating Its Parameters Using the Method of L-Moments. Paper presented at *RELIK – Reprodukce lidského kapitálu* held on 05–06 December 2011, Prague, CD. ISBN 978-80-86175-75-1.

BÍLKOVÁ, D. Use of the L-Moments Method in Modeling the Wage Distribution. Paper presented at *Aplimat* held on 01–04 February 2011, Bratislava, pp. 1471–1481. ISBN 978-80-89313-52-5.

BÍLKOVÁ, D. MALÁ, I. Application of the L-Moment Method when Modelling the Income Distribution in the Czech Republic. *Austrian Journal of Statistics,* 2012, Vol. 41, No. 2, pp. 125–132. ISSN 1026-597X.

ELAMIR, E. A. H., SEHEULT, A. H. Trimmed L-moments. *Computational Statistics & Data Analysis*, 2003, Vol. 43, No. 3, pp. 299–314. ISSN 0167-9473.

HÁTLE, J., HUSTOPECKÝ, J. NOVÁK, I. *Modelování a krátkodobá předpověď příjmových rozdělení – Výzkumná práce č. 66* (Modeling and Short-Term Forecast of Income Distribution – Research Paper No. 66). Prague: Výzkumný ústav sociálně ekonomických informací (Researchn Institute for Social and Economic Information), 1975.

HOSKING, J. R. M. L-moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *Journal of the Royal Statistical Society* (Series B), 1990, Vol. 52, No. 1, pp. 105–124. ISSN 1467-9868.

KYSELÝ, J., PICEK, J. Regional Growth Curves and Improved Design Value Estimates of Extreme Precipitation Events in the Czech Republic. *Climate research*, 2007, Vol. 33, pp. 243–255. ISSN 0936-577X.

MAREK, L. Wage development in the Czech Republic for the past 16 years. Paper presented at *AIESA – Budovanie společnosti založenej na vedomostiach* held on 10–11 November 2011, Bratislava, pp. 1–7. ISBN 978-80-225-3312-6.

MAREK, L., VRABEC, M. Mixture Normal Density Functions as a Model Wage Distribution. Paper presented at *Economic, Marketing and Management* held on 19–20 January 2013, Dubai, pp. 69–74. ISBN 978-981-07-5039-8. ISSN 2010-4626.
MAREK, L., VRABEC, M. Model wage distribution – mixture density functions. *Internatrional Journal of Economics and Statistics* [online], 2013, Vol. 1, No. 3, pp. 113–121. ISSN 2309-0685. <http://www.naun.org/main/NAUN/economics/c052015-092.pdf>.
MAREK, L., VRABEC, M. Wage Distributions Models. Paper presented at EBA 2013 – *Economics and Business Administration* held on 16–19 July 2013, Greece, Rhodes Island, pp. 180–183. ISBN 978-1-61804-198-2.
SERFLING, R. J. *Approximation Theorems of Mathematical Statistics*. New York: John Wiley & Sons, 1980, 375 p. ISBN 0-471-21927-4.
ULRYCH, T. J., VELIS, D. R., WOODBURY, A. D., SACCHI, M. D. L-moments and C-moments. *Stoch. Environ. Res. Risk Asses*, 2000, Vol. 14, pp. 50–68. ISSN 1436-3240.

Recent Publications and Events

New Publications of the Czech Statistical Office

Czech Republic in International Comparison 2013. Selected Indicators. Prague: CZSO, 2014.

Dlouhodobý vývoj bytové výstavby v České republice (Long-term development of housing construction in the Czech Republic). *Prague: CZSO, 2013.*

DUBSKÁ, D. *Hmotné bohatství v České republice: posiluje ho vývoj na realitním trhu?* (Material wealth in the Czech Republic: is it strengthened by the development on the real estate market?). Prague: CZSO, 2014.

DUBSKÁ, D., KAMENICKÝ, J., KOHOUTOVÁ, I., KUČERA, L. *Tendencies and Factors of Macroeconomic Development and Quality of Life in the Czech Republic in 2012. Prague: CZSO, 2013.*

DUBSKÁ, D., KAMENICKÝ, J., KUČERA, L. *The Czech Economy Development in 2013. Prague: CZSO, 2014.*

Green Growth in the Czech Republic. Selected Indicators 2013. Prague: CZSO, 2014.

Inovační aktivity podniků v České republice v letech 2010–2012 (Innovation activities of enterprises in the Czech Republic in 2010–2012). Prague: CZSO, 2014.

Náboženská víra obyvatel podle výsledků sčítání lidu 2011 (Religious Belief of the Population According to Results of the Census 2011). Prague: CZSO, 2014.

Sčítání lidu, domů a bytů 2011 (Population and Housing Census 2011). Prague: CZSO, 2013.

Spotřeba potravin 1948–2012 (Food Consumption 1948–2012). Prague: CZSO, 2013.

Ukazatele výzkumu a vývoje za rok 2012 (Research and Development Indicators for 2012). Prague: CZSO, 2013.

Other Selected Publications

DOLEJŠ, J. *Ekonomická statistika* (Economic Statistics). Hradec Králové: Gaudeamus, 2013.

HUFF, D. *Jak lhát se statistikou* (How to Lie with Statistics). Prague: Brána, 2013.

Konkurenční schopnost České republiky 2011–2012 (Competitiveness of the Czech Republic 2011–2012). Prague: CES VŠEM, 2013.

MARTIN, V., HURN, S., HARRIS, D. *Econometric Modelling with Time Series.* Cambridge University Press, 2013.

MELOUN, M., MILITKÝ, J. *Interaktivní statistická analýza dat* (Interactive Statistical Data Analysis). Prague: Karolinum, 2012.

MINAŘÍK, B., BORŮVKOVÁ, J., VYSTRČIL, M. *Analýzy v regionálním rozvoji (Analyses in regional development).* Prague: Professional Publishing, 2013.

Sustainable development in the European Union. Luxembourg: Eurostat, 2013.

TIJMS, H. *Understanding Probability.* Cambridge University Press, 2013.

WILLINK, R. *Measurement Uncertainty and Probability.* Cambridge University Press, 2013.

ZEMAN, K. *Vývoj vlastnictví k půdě a souvisejících procesů na území ČR od roku 1918 do současné doby* (Development of the Land Ownership and Related Processes in the Czech Republic from 1918 to the Present Time). Prague: University of Economics, 2013.

Conferences

The ***12th Global Forum on Tourism Statistics GFTS 2014*** took place during ***15–16 May 2014 in Prague, Czech Republic.*** The event was organised jointly by the Ministry of Regional Development of the Czech Republic, the Czech Statistical Office, the Czech Tourism Board, the Statistical Office of the European Union (EUROSTAT) and the Organisation for Economic Cooperation and Development (OECD). The aim was to discuss major technical issues concerning the establishment of harmonised tourism statistics in an environment that strengthens co-operation among governments, the private sector, researchers, academics, OECD and EU member and non-member countries and international organisations. More information available at: *http://www.tsf2014prague.cz*.

The ***21st International Conference on Computational Statistics COMPSTAT 2014*** will be held ***from 19th to 22nd August 2014*** at the International Conference Centre in ***Geneva, Switzerland***. The conference aims at bringing together researchers and practitioners to discuss recent developments in computational methods, methodologies for data analysis and applications in statistics. More information available at: *http://compstat2014.org.*

The ***17th International Scientific Conference Applications of Mathematics and Statistics in Economics AMSE 2014*** will take place ***from 27th to 31st August 2014*** in ***Jerzmanowice, Poland***. The purpose of the conference is to acquaint the participants of the conference with the latest mathematical and statistical methods that can be used in solving theoretical and practical economic problems. AMSE 2014 is organized by the University of Economics, Prague, Czech Republic (Faculty of Informatics and Statistics, Department of Statistics and Probability), Matej Bel University, Banská Bystrica, Slovakia (Faculty of Economics, Department of Quantitative Methods and Information Technology) and Wroclaw University of Economics, Wroclaw, Poland (Department of Statistics). More information available at: *www.amse.ue.wroc.pl.*

The ***8th International Days of Statistics and Economics MSED 2014*** will be held during ***11–13 September 2014*** at the University of Economics, ***Prague, Czech Republic***, organized by the Department of Statistics and Probability and the Department of Microeconomics (University of Economics, Prague, Czech Republic), the Faculty of Business Economics (University of Economics with seat in Košice, Slovakia) and the ESC Rennes International School of Business (France). The aim of the conference is to present and discuss current problems of Statistics, Demography, Economics and Management and their mutual interconnection. More information available at: *msed.vse.cz.*

The ***60th World Statistics Congress ISI 2015*** will take place during ***26–31 July 2015*** in ***Rio de Janeiro, Brazil***. The congress will bring together members of the statistical community to present, discuss, promote and disseminate research and best practice in every field of Statistics and its applications. More information available at: *http://www.isi2015.ibge.gov.br.*

Papers
We publish articles focused at theoretical and applied statistics, mathematical and statistical methods, conception of official (state) statistics, statistical education, applied economics and econometrics, economic, social and environmental analyses, economic indicators, social and environmental issues in terms of statistics or economics, and regional development issues.

The journal of *Statistika* has the following sections:
The *Analyses* section publishes high quality, complex, and advanced analyses based on the official statistics data focused on economic, environmental, and social spheres. Papers shall have up to 12,000 words or up to twenty (20) 1.5-spaced pages.

The *Methodology* section gives space for the discussion on potential approaches to the statistical description of social, economic, and environmental phenomena, development of indicators, estimation issues, etc. Papers shall have up to 12,000 words or up to twenty (20) 1.5-spaced pages.

The *Book Review* section brings reviews of recent books in the field of the official statistics. Reviews shall have up to 600 words or one (1) 1.5-spaced page.

In the *Information* section we publish informative (descriptive) texts. The maximum range of information is 6 000 words or up to 10 1.5-spaced pages.

Language
The submission language is English only. Authors are expected to refer to a native language speaker in case they are not sure of language quality of their papers.

Recommended Paper Structure
Title (e.g. On Laconic and Informative Titles) — Authors and Contacts — Abstract (max. 160 words) — Keywords (max. 6 words / phrases) — JEL classification code — Introduction — … — Conclusion — Annex — Acknowledgments — References — Tables and Figures

Authors and Contacts
Rudolf Novak*, Institution Name, Street, City, Country
Jonathan Davis, Institution Name, Street, City, Country
* Corresponding author: e-mail: rudolf.novak@domain-name.cz, phone: (+420) 111 222 333

Main Text Format
Times 12 (main text), 1.5 spacing between lines. Page numbers in the lower right-hand corner. *Italics* can be used in the text if necessary. *Do not* use **bold** or underline in the text. Paper parts numbering: 1, 1.1, 1.2, etc.

Headings
1 FIRST-LEVEL HEADING (Times New Roman 12, bold)
1.1 Second-level heading (Times New Roman 12, bold)
1.1.1 Third-level heading (Times New Roman 12, bold italic)

Footnotes
Footnotes should be used sparingly. Do not use endnotes. Do not use footnotes for citing references (except headings).

References in the Text
Place reference in the text enclosing authors' names and the year of the reference, e.g. "White (2009) points out that…", "… recent literature (Atkinson et Black, 2010a, 2010b, 2011, Chase et al., 2011, pp. 12–14) conclude…". Note the use of alphabetical order. Include page numbers if appropriate.

List of References
Arrange list of references alphabetically. Use the following reference styles: [for a book] HICKS, J. *Value and Capital: An inquiry into some fundamental principles of economic theory.* Oxford: Clarendon Press, 1939. [for chapter in an edited book] DASGUPTA, P. et al. Intergenerational Equity, Social Discount Rates and Global Warming. In PORTNEY, P., WEYANT, J., eds. *Discounting and Intergenerational Equity.* Washington, D.C.: Resources for the Future, 1999. [for a journal] HRONOVÁ, S., HINDLS, R., ČABLA, A. Conjunctural Evolution of the Czech Economy. *Statistika, Economy and Statistics Journal,* 2011, 3 (September), pp. 4–17. [for an online source] CZECH COAL. *Annual Report and Financial Statement 2007* [online]. Prague: Czech Coal, 2008. [cit. 20.9.2008]. <http://www.czechcoal.cz/cs/ur/zprava/ur2007cz.pdf>.

Tables
Provide each table on a separate page. Indicate position of the table by placing in the text "insert Table 1 about here". Number tables in the order of appearance Table 1, Table 2, etc. Each table should be titled (e.g. Table 1 Self-explanatory title). Refer to tables using their numbers (e.g. see Table 1, Table A1 in the Annex). Try to break one large table into several smaller tables, whenever possible. Separate thousands with a *space* (e.g. 1 528 000) and decimal points with a *dot* (e.g. 1.0). Specify the data source below the tables.

Figures
Figure is any graphical object other than table. Attach each figure as a separate file. Indicate position of the figure by placing in the text "insert Figure 1 about here". Number figures in the order of appearance Figure 1, Figure 2, etc. Each figure should be titled (e.g. Figure 1 Self-explanatory title). Refer to figures using their numbers (e.g. see Figure 1, Figure A1 in the Annex).

Figures should be accompanied by the *.xls, *.xlsx table with the source data. Please provide cartograms in the vector format. Other graphic objects should be provided in *.tif, *.jpg, *.eps formats. Do not supply low-resolution files optimized for the screen use. Specify the source below the figures.

Formulas
Formulas should be prepared in formula editor in the same text format (Times 12) as the main text.

Paper Submission
Please email your papers in *.doc, *.docx or *.pdf formats to statistika.journal@czso.cz. All papers are subject to double-blind peer review procedure. You will be informed by our managing editor about all necessary details and terms.

Contacts
Journal of Statistika | Czech Statistical Office
Na padesátém 81 | 100 82 Prague 10 | Czech Republic
e-mail: statistika.journal@czso.cz
web: www.czso.cz/statistika_journal

Managing Editor: Jiří Novotný
phone: (+420) 274 052 017
fax: (+420) 274 052 133
e-mail: statistika.journal@czso.cz
web: www.czso.cz/statistika_journal
address: Czech Statistical Office | Na padesátém 81 | 100 82 Prague 10 | Czech Republic

Subscription price (4 issues yearly)
CZK 372 (incl. postage) for the Czech Republic,
EUR 110 or USD 165 (incl. postage) for other countries.
Printed copies can be bought at the Publications Shop of the Czech Statistical Office (CZK 66 per copy).
address: Na padesátém 81 | 100 82 Prague 10 | Czech Republic

Subscriptions and orders
MYRIS TRADE, s. r. o.
P. O. BOX 2 | 142 01 Prague 4 | Czech Republic
phone: (+420) 234 035 200,
fax: (+420) 234 035 207
e-mail: myris@myris.cz

Design: Toman Design
Layout: Ondřej Pazdera
Typesetting: Josef Neckář
Print: Czech Statistical Office

94th year of the series of professional economy and statistics journals of the State Statistical Service in the Czech Republic: **Statistika** (since 1964), **Statistika a kontrola** (1962 – 1963), **Statistický obzor** (1931 – 1961), **Československý statistický věstník** (1920 – 1930).

Published by the Czech Statistical Office
ISSN 1804-8765 (Online)
ISSN 0322-788X (Print)
Reg. MK CR E 4684