STATISTIKA

STATISTICS
AND ECONOMY
JOURNAL

VOL. **100** (1) 2020

CZECH
STATISTICAL
OFFICE

CONTENTS

ANALYSES

CONSULTATION

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 professional double-blind peer reviewed open access journal included in the citation database of peer-reviewed literature **Scopus** (since 2015), in the **Web of Science** *Emerging Sources Citation Index* (since 2016), and also in other international databases of scientific journals. 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's 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

Economic Growth, Business Cycles and Okun's Law: Unobserved Components Approach

Andrea Čížků [1] | *University of Economics, Prague, Czech Republic*

Abstract

Clark's (1989) bivariate unobserved components model is applied in order to estimate and analyse the trend and cycle of GDP and the unemployment rate as well as to quantify and discuss the relationship known as Okun's law. Empirical analysis is performed for 28 European countries for a time period including the current economic crisis – the end period is 2018 Q4 for all economies and the beginning period ranges from 1983 Q1 to 2000 Q1 according to data availability. Important results indicate that in virtually all European countries: (1) the growth of the trend component of GDP decreased systematically after the crisis; (2) the output gap improved in the last five years – this finding proved to be quite robust as it was also confirmed by Hodrick-Prescott estimates of the output gap for different smoothing parameter values; (3) the trend component of the unemployment rate turned out to be constant over time, indicating that possible hysteresis effects have not played an important role in European labour markets; (4) the output gap and cyclical unemployment rate are highly negatively correlated, confirming the strength and validity of Okun's law across European countries.

Keywords	***JEL code***
Unobserved components, Clark's bivariate model, Okun's law, Kalman filter, HP filter	*C32, E24, E32*

INTRODUCTION

Economists concern themselves with unobserved features of an economy such as potential (trend) output or output gap (cycle) in order to separate long-run and short-run components of economic variables. This is usually motivated by the goal of dampening economic fluctuations by keeping output at its potential level. Trend-cycle decomposition also enables the assessment of long-term effects of the current economic crisis, which is a widely discussed issue in economic literature these days. Ball (2014) found empirical evidence supporting the hypothesis of permanent effects of deep recessions on output. Similar findings are reported by Barro (2001), Cerra and Saxena (2005, 2008) and Haltmaier (2012).

[1] University of Economics, Prague, Faculty of Informatics and Statistics, Department of Econometrics, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: andrea.cizku@vse.cz, phone: (+420)224095443. ORCID: 0000-0003-3053-3222.

Trend-cycle decomposition is usually performed in empirical literature for aggregated European data, such as by Azevedo et al. (2003), Berger and Everaert (2008), Berger (2011), Bernhofer et al. (2014), Chen and Mills (2012), Galati et al. (2016), Lemoine et al. (2010), Orlandi and Pichelmann (2000) and Proietti (2004).

It is also common for authors to decompose variables for only one specific economy or a small group of countries, e.g. Hájek and Bezděk (2000), Hájková and Hurník (2007), Kloudová (2013), and Beneš and N'Diaye (2004) for the Czech economy; Bernardi and Di Ruggiero (2014) for Italy; Boďa et al. (2015), and Boďa and Považanová (2019) for Visegrad group countries; Jemec (2012) for Slovenia; Volos and Hadjixenophontos (2015) for Cyprus; and Ochotnický (2008), and Zimková and Barochovský (2007) for Slovakia.

This paper contributes to existing literature by providing an extensive empirical investigation and comparison of 28 European economies taking into account their heterogeneity. The focus is placed on: (1) characterizing changes in the trend and cycle of GDP and the unemployment rate during the current crisis, which enables the assessment of the long-run and transient influence of the crisis on these two variables; (2) estimation and discussion of the relationship between cyclical components of GDP and the unemployment rate, known as Okun's law; (3) making a comparison of these features across individual European countries.

Primarily, the trend-cycle decomposition is performed by applying the unobserved components (UC) methodology. However, Hodrick and Prescott's (1997) (HP) filter with different smoothing parameter values is applied as a robustness check as well.

Such a vast empirical study has only been performed by Ball (2014), Ball et al. (2017), and Brůha and Polanský (2015). Nonetheless, these papers are methodologically based on simple methods such as the production function approach, HP filter or analysis of correlations.

There are various methods for decomposing economic variables. A survey of alternative methodologies was performed by Dupasquier et al. (1999). This paper applies structural time series modelling as advocated by Harvey (1989), which uses the state-space form and Kalman filtering in order to estimate unobserved components of time series. Specifically, Clark's (1989) unobserved components (UC) model is employed in order to explicitly model the trend and cycle of GDP and the unemployment rate. Harvey and Jaeger (1993) showed that this technique of detrending economic time series is superior to ARIMA modelling and to mechanical statistical tools like the HP filter.

The structure of the paper is as follows. The model is formulated in Section 1 and data are described in Section 2. Econometric estimation of parameters is discussed in Section 3, and the unobserved trend and cycle of GDP and the unemployment rate are presented in Section 4. Subsequently, Section 5 compares the calculated output gap with that obtained using the HP filter. The final section summarizes the main findings and concludes.

1 MODEL

A minor modification of Clark's (1989) bivariate UC model as specified by Kim and Nelson (1999) is applied in this paper in order to study GDP and unemployment in European countries. This model extends the univariate model of GDP decomposition formulated and estimated for the US economy by Clark (1987), which is given as follows:

$$y_t = n_t + x_t, \tag{1}$$

$$n_t = g_{t-1} + n_{t-1} + v_t,\ v_t \sim i.i.d.\ N\left(0, \sigma_v^2\right), \tag{2}$$

$$g_t = g_{t-1} + w_t,\ w_t \sim i.i.d.\ N\left(0, \sigma_w^2\right), \tag{3}$$

$$x_t = \phi_1 \cdot x_{t-1} + \phi_2 \cdot x_{t-2} + e_t,\ e_t \sim i.i.d.\ N\left(0, \sigma_e^2\right), \tag{4}$$

where y_t is the log of real GDP, n_t represents the stochastic trend component evolving according to a random walk process with drift g_t, which is itself the random walk, and x_t is a stationary cyclical component (output gap). Second-order autoregression for x_t allows the output gap to exhibit cyclical movements. It is also possible to derive AR(2) specification for the gap from a standard model of cyclical components based on sine and cosine functions (Orlandi and Pichelmann, 2000).

Clark's (1987) formulated univariate UC model given by Formulas (1)–(4) was extended by Clark (1989) and later also by Kim and Nelson (1999) by including the unemployment rate U_t in the following fashion:[2]

$$U_t = L_t + C_t, \tag{5}$$

$$L_t = L_{t-1} + \varepsilon_t,\ \varepsilon_t \sim i.i.d.\ N\left(0, \sigma_\varepsilon^2\right), \tag{6}$$

$$C_t = \alpha_0 \cdot x_t + \alpha_1 \cdot x_{t-1} + \eta_t,\ \eta_t \sim i.i.d.\ N\left(0, \sigma_\eta^2\right), \tag{7}$$

where L_t is the trend of the unemployment rate, C_t represents the cyclical stationary component of the unemployment rate, and ε_t and η_t are white noise processes. Random errors v_t, w_t, e_t, ε_t and η_t in the formulated model (1)–(7) are all not only assumed serially but also mutually independent.

Formula (7) is interpreted as Okun's law postulating a negative relation between the output gap and the unemployment rate gap. It corresponds to Clark's (1989) formulation in that it is assumed that only the current output gap x_t and the output gap lagged one period x_{t-1} have an influence on the cyclical unemployment rate C_t. The same specification was also applied by Berger (2011). Kim and Nelson (1999) also included the output gap lagged two periods x_{t-2} in Okun's law reflecting that the unemployment rate is a lagging indicator of the business cycle. Nonetheless, this specification turned out to be unsuitable for many European countries as it caused higher standard errors of many estimated parameters (not only relating to Okun's law).

Firstly, it is worth mentioning that output and unemployment rate are modelled simultaneously, not separately. The output decomposition model (1)-(4) is interconnected with the unemployment rate model (5)-(7) by Okun's law in the form of Formula (7). Estimation of unobserved cyclical components of the output and unemployment rate therefore utilizes this structural relationship between these two variables. Secondly, the applied specification of Okun's law (7) postulates that the output gap influences the cyclical unemployment rate so that assumed causality goes from output to unemployment, which is a lagging indicator. This is a slight modification of Okun's (1962) empirical investigation where he never postulated this one-way causality.

2 DATA

Seasonally adjusted quarterly GDP and unemployment rate data for 28 European economies were obtained from Eurostat (2018). GDP is measured in chain-linked volumes (reference year 2010, million euro) and the name of the series in the Eurostat database is "GDP and main components (output, expenditure and income) [namq_10_gdp]". Unemployment rate includes all ages of unemployed workers and relates

[2] Clark's (1989) extension slightly differs from that performed by Kim and Nelson (1999). Clark (1989) specifies the GDP trend component as a random walk without a drift term . Kim and Nelson (1999) follow Clark's (1987) original specification (1)-(4) in their extension to the bivariate model, which is an approach taken in this paper as well.

the number of unemployed persons to the active population (expressed as a number between 0 and 1, not in per cent), and the title of the source data in Eurostat is "Unemployment by sex and age – quarterly average [une_rt_q]". The observable variable y_t is represented by the natural logarithm of the GDP, and the variable U_t corresponds directly to the unemployment rate.

The last observation in the data set is 2018 Q4 for all 28 countries. However, the first observation differs in individual European economies due to data availability and starting period of the data for countries ranges from 1983 Q1 to 2000 Q1. The starting date for each economy is indicated in the following Table 1, which contains estimation results (placed in Section 3) in the first column below the country name.

3 ECONOMETRIC ESTIMATION AND DISCUSSION

The bivariate UC model (1)–(7) was expressed in state-space form and econometrically estimated by maximizing the likelihood function, which was performed numerically in MATLAB. The likelihood function was constructed by applying the Kalman filter algorithm, which was implemented in its square root version in order to achieve better numerical precision. Kalman filtering was initialized by standard method – unconditional mean and variance were used as initial values for stationary cyclical components and a diffuse prior was applied for non-stationary trend components. Parameter transformations ensuring non-negativity of standard deviations of random errors and stationarity of autoregressive process (4) were also performed.

Results for 28 European countries are summarized in the following Table 1. Standard errors of estimated coefficients are shown in parentheses below estimated parameters. For readers' convenience, the symbols ***, ** and * are used as well in order to indicate statistical significance at the 1%, 5% and 10% level, respectively. Standard deviations of random shocks are multiplied by 100 in order to obtain more readable results. Alternatively, data for the log of GDP and unemployment rate could be multiplied by 100 in order to increase estimated standard errors of random shocks 100 times.

Table 1 Estimation results of the bivariate UC model (1)–(7) for European countries

Country (start date)[3]	GDP					Unemployment rate			
	$\hat{\phi}_1$	$\hat{\phi}_2$	$100\hat{\sigma}_e$	$100\hat{\sigma}_w$	$100\hat{\sigma}_v$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$100\hat{\sigma}_\varepsilon$	$100\hat{\sigma}_\eta$
Austria	1.53***	−0.58***	0.37***	0.04	0.31***	0.05	−0.34***	0.16***	0.09***
(1996 Q1)	(0.13)	(0.13)	(0.09)	(0.03)	(0.09)	(0.09)	(0.12)	(0.04)	(0.03)
Belgium	1.46***	−0.54***	0.39***	0.03	0.00	−0.05	−0.22**	0.29***	0.18***
(1995 Q1)	(0.14)	(0.14)	(0.09)	(0.02)	(5.30)	(0.10)	(0.10)	(0.04)	(0.04)
Bulgaria	1.77***	−0.78***	0.21**	0.13**	0.76***	−0.11	−1.11	0.34**	0.01
(2000 Q1)	(0.09)	(0.09)	(0.08)	(0.06)	(0.07)	(0.78)	(0.88)	(1.16)	(0.16)
Croatia	1.79***	−0.80***	0.24**	0.11	1.13***	−0.52	−0.42	0.02	0.09
(2000 Q1)	(0.08)	(0.09)	(0.10)	(0.07)	(0.10)	(0.59)	(0.52)	(0.90)	(0.07)
Cyprus	1.81***	−0.82***	0.35***	0.09*	0.67***	−0.32	−0.38	0.34***	0.15
(2000 Q1)	(0.09)	(0.09)	(0.11)	(0.05)	(0.08)	(0.32)	(0.31)	(0.12)	(0.10)
Czech Rep.	1.76***	−0.77***	0.38***	0.28***	0.30***	−0.18*	−0.36***	0.09	0.07
(1996 Q1)	(0.09)	(0.09)	(0.08)	(0.06)	(0.06)	(0.09)	(0.12)	(0.11)	(0.05)
Denmark	1.68***	−0.71***	0.31***	0.00	0.76***	0.17	−0.68**	0.00	0.07
(1995 Q1)	(0.11)	(0.11)	(0.08)	(0.11)	(0.07)	(0.24)	(0.29)	(2.91)	(0.05)
Estonia	1.66***	−0.69***	0.94***	0.09	1.51***	−0.13	−0.42**	0.42	0.37***
(2000 Q1)	(0.12)	(0.12)	(0.25)	(0.08)	(0.19)	(0.18)	(0.19)	(0.28)	(0.13)
Finland	1.82***	−0.83***	0.40***	0.05	1.04***	0.03	−0.62***	0.05	0.00
(1990 Q1)	(0.06)	(0.06)	(0.08)	(0.04)	(0.08)	(0.18)	(0.22)	(0.25)	(0.63)
France	1.78***	−0.79***	0.20***	0.04**	0.29***	−0.01	−0.45***	0.10***	0.00
(1983 Q1)	(0.06)	(0.06)	(0.03)	(0.02)	(0.02)	(0.10)	(0.12)	(0.02)	(0.16)
Germany	1.70***	−0.71***	0.33***	0.09**	0.64***	−0.18**	−0.25***	0.06	0.00
(1991 Q1)	(0.09)	(0.09)	(0.06)	(0.04)	(0.05)	(0.08)	(0.09)	(0.06)	(0.11)

Table 1 (continuation)

Country (start date)[3]	GDP $\hat{\phi}_1$	$\hat{\phi}_2$	$100\hat{\sigma}_e$	$100\hat{\sigma}_w$	$100\hat{\sigma}_v$	Unemployment rate $\hat{\alpha}_0$	$\hat{\alpha}_1$	$100\hat{\sigma}_\varepsilon$	$100\hat{\sigma}_\eta$
Greece	1.85***	−0.86***	0.37***	0.12*	1.11***	−0.87*	0.07	0.01	0.11
(1998 Q2)	(0.07)	(0.07)	(0.10)	(0.07)	(0.10)	(0.47)	(0.39)	(2.41)	(0.09
Hungary	1.78***	−0.80***	0.11	0.22***	0.67***	−0.41	−1.08	0.20**	0.01
(1996 Q1)	(0.12)	(0.12)	(0.08)	(0.06)	(0.06)	(1.62)	(1.97)	(0.10)	(0.41)
Ireland	1.80***	−0.81***	0.19	0.00	2.82***	−1.34	−0.01	0.00	0.15**
(1995 Q1)	(0.09)	(0.09)	(0.17)	(0.57)	(0.21)	(3.40)	(2.83)	(10.87)	(0.08)
Italy	1.78***	−0.80***	0.31***	0.00	0.45***	−0.08	−0.28**	0.20***	0.05
(1995 Q1)	(0.09)	(0.09)	(0.07)	(0.11)	(0.05)	(0.14)	(0.14)	(0.04)	(0.06)
Latvia	1.69***	−0.71***	0.83***	0.21**	1.19***	−0.38**	−0.23	0.19	0.37***
(1998 Q2)	(0.10)	(0.10)	(0.20)	(0.09)	(0.15)	(0.17)	(0.14)	(0.46)	(0.09)
Lithuania	1.74***	−0.75***	0.73***	0.07	1.61***	−0.41*	−0.23	0.01	0.28***
(1998 Q1)	(0.09)	(0.09)	(0.17)	(0.08)	(0.15)	(0.22)	(0.19)	(8.00)	(0.09)
Luxembourg	1.64***	−0.66***	0.21	0.05	1.61***	−0.65	0.03	0.01	0.03
(1995 Q1)	(0.19)	(0.19)	(0.15)	(0.07)	(0.12)	(0.88)	(0.58)	(0.79)	(0.11)
Malta	0.45	0.53*	0.82*	0.13*	1.05***	−0.03	−0.07	0.33***	0.09
(2000 Q1)	(0.32)	(0.29)	(0.49)	(0.08)	(0.37)	(0.10)	(0.11)	(0.05)	(0.08)
Netherlands	1.86***	−0.86***	0.19***	0.03	0.50***	0.17	−0.69***	0.00	0.03
(1996 Q1)	(0.06)	(0.06)	(0.04)	(0.03)	(0.04)	(0.23)	(0.26)	(0.97)	(0.04)
Norway	1.67***	−0.69***	0.13	0.03	1.16***	0.39	−1.20	0.07	0.00
(1989 Q1)	(0.14)	(0.15)	(0.09)	(0.04)	(0.08)	(0.65)	(1.05)	(0.14)	(0.48)
Poland	1.78***	−0.79***	0.11*	0.00	0.95***	−3.40	0.57	0.00	0.00
(1997 Q1)	(0.10)	(0.10)	(0.08)	(0.18)	(0.07)	(4.58)	(3.32)	(0.20)	(0.02)
Portugal	1.76***	−0.77***	0.31***	0.07**	0.52***	−0.23	−0.54**	0.17	0.10
(1995 Q1)	(0.09)	(0.09)	(0.07)	(0.03)	(0.05)	(0.19)	(0.21)	(0.11)	(0.06)
Romania	0.90**	0.06	0.30	0.25**	1.57***	−0.96	−0.36	0.00	0.00
(1997 Q1)	(0.37)	(0.32)	(0.36)	(0.11)	(0.15)	(2.26)	(0.79)	(31.19)	(3.29)
Slovenia	1.75***	−0.76***	0.52***	0.07	0.71***	0.04	−0.31**	0.25***	0.14**
(1996 Q1)	(0.10)	(0.11)	(0.13)	(0.05)	(0.10)	(0.12)	(0.13)	(0.06)	(0.05)
Spain	1.88***	−0.88***	0.21***	0.09***	0.13***	−1.12***	0.01	0.28***	0.00
(1995 Q1)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.30)	(0.23)	(0.07)	(1.06)
Sweden	1.61***	−0.65***	0.34***	0.02	0.72***	0.05	−0.55***	0.00	0.08**
(1993 Q1)	(0.07)	(0.08)	(0.07)	(0.03)	(0.06)	(0.12)	(0.16)	(1.86)	(0.04)
UK	1.82***	−0.83***	0.24***	0.05**	0.44***	−0.11	−0.43***	0.10***	0.00
(1983 Q1)	(0.06)	(0.06)	(0.04)	(0.02	(0.03)	(0.11)	(0.13)	(0.04)	(0.21)

Source: Authors own calculations based on data obtained from Eurostat (2018)

Table 1 contains detailed estimation results for all 28 studied economies, which enables differences between individual European countries to be assessed. Nonetheless, some sort of aggregation by using median values to characterize a typical European economy will often be utilized in subsequent discussion for two reasons. Firstly, it facilitates the reader's basic orientation in these extensive results. Secondly, it enables a comparison with findings of other empirical studies analysing aggregated European data.

3.1 Parameters relating to GDP decomposition

3.1.1 Autoregressive parameters ϕ_1 and ϕ_2

Parameters ϕ_1 and ϕ_2 are statistically significant at the 1% level of significance with the only exceptions being Malta and Romania. The estimated parameters $\hat{\phi}_1 = 0.90$ and $\hat{\phi}_2 = 0.06$ for Romania resemble

[3] The last observation in the sample is 2018 Q4 for every country.

closely a stationary AR(1) process. Estimates for Malta are quite atypical, which applies not only for parameters ϕ_1 and ϕ_2.

Another substantially robust finding is that the sum $\hat{\phi}_1 + \hat{\phi}_2$ is close to one (except for Malta and Romania). This means that the output gap x_t is highly persistent. Clark (1989) reports the following results:

a) $\hat{\phi}_1 = 1.47$, $\hat{\phi}_2 = -0.59$ for the US economy (1947 Q1–1986 Q2),

b) $\hat{\phi}_1 = 1.59$, $\hat{\phi}_2 = -0.63$ for Canada (1955 Q1–1986 Q2),

c) $\hat{\phi}_1 = 1.73$, $\hat{\phi}_2 = -0.73$ for West Germany (1960 Q1–1986 Q2),

d) $\hat{\phi}_1 = 1.82$, $\hat{\phi}_2 = -0.82$ for the United Kingdom (1960 Q1–1986 Q2),

e) $\hat{\phi}_1 = 0.54$, $\hat{\phi}_2 = 0.20$ for Japan (1960 Q1–1986 Q2).

The sum of these parameters estimated for the US economy, $\hat{\phi}_1 + \hat{\phi}_2 = 0.88$, is not as close to 1 as in most European countries. This is also confirmed by Kim and Nelson's (1999) estimates for the US economy (1952 Q1–1995 Q3): $\hat{\phi}_1 = 1.44$, $\hat{\phi}_2 = -0.52$. When excluding Malta and Romania, the median value of $\hat{\phi}_1 + \hat{\phi}_2$ from Table 1 is 0.99. The output gap in virtually all European countries is thus remarkably more persistent than the corresponding gap in the US economy. It is also confirmed that estimates presented in Table 1 for Germany ($\hat{\phi}_1 = 1.70$, $\hat{\phi}_2 = -0.71$) and for the United Kingdom ($\hat{\phi}_1 = 1.82$, $\hat{\phi}_2 = -0.83$) are very close to those reported by Clark (1989).

Similar results regarding the persistence of Visegrad group countries were found by Boďa and Považanová (2019). Berger's (2011) estimates for aggregated European data from 1970 to 2005 are also in line with these conclusions as he estimated these parameters as $\hat{\phi}_1 = 1.88$ and $\hat{\phi}_2 = -0.91$ (with standard errors of 0.05 and 0.04, respectively).

These findings regarding persistence are also supported by Chen and Mills (2012), who analysed aggregate euro area data ranging from 1970 Q1 to 2009 Q2. Nonetheless, some comments are needed before we report their results as they did not model the cyclical component of output directly as the AR(2) process. Instead, they modelled the cyclical component (ψ_t in their notation) using sine and cosine functions as follows:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \tag{8}$$

where κ_t and κ_t^* are mutually and serially independent random errors with $Var(\kappa_t) = Var(\kappa_t^*) = \sigma_\kappa^2$, $0 \le \rho < 1$ is the damping parameter and λ_c is the cycle frequency (in radians). Orlandi and Pichelmann (2000) showed that specification (8) is equivalent to the following AR(2) model:

$$\psi_t = (2\rho\cos\lambda_c)\psi_{t-1} - \rho^2\psi_{t-1} + \zeta_t, \tag{9}$$

where random error ζ_t is given by:

$$\zeta_t = (1 - \rho\cos\lambda_c L)\kappa_t - (\rho\sin\lambda_c L)\kappa_t^*, \tag{10}$$

where L is the lag operator.

Formula (9) is already comparable to specification (4). Chen and Mills (2012) estimate parameter ρ and the period of the cycle $2\pi / \lambda_c$ for aggregate euro area data as $\hat{\rho} = 0.82$ and $2\pi / \hat{\lambda}_c = 35.34$. This implies that $\hat{\lambda}_c = 0.18$, $\hat{\phi}_1 = 2\hat{\rho}\cos\hat{\lambda}_c = 1.61$ and $\hat{\phi}_2 = -\hat{\rho}^2 = -0.67$. This result is somewhere between Clark's (1989) estimate for the US economy and Berger's (2011) estimate for Europe.

3.1.2 Volatility of output gap σ_e

The random error e_t in Formula (4) describing the development of the output gap plays an important role as the coefficient σ_e is statistically significant in the majority of studied countries. This parameter failed to be statistically significant only for Hungary, Ireland, Luxembourg, Norway and Romania. The output gap is thus stable in these economies.

Chen and Mills (2012) estimated that $100\hat{\sigma}_\kappa = 0.27$ and it follows from that:

$$
\begin{aligned}
Var(\zeta_t) &= \sigma_\kappa^2 + \rho^2 \cos^2 \lambda_c \sigma_\kappa^2 + \rho^2 \sin^2 \lambda_c \sigma_\kappa^2, \\
&= \sigma_\kappa^2 (1 + \rho^2).
\end{aligned}
$$

Correspondence between (9) and (4) then yields $\sigma_e^2 = Var(\zeta_t)$. Therefore, $100\hat{\sigma}_e = 100\hat{\sigma}_\kappa \cdot \sqrt{1+\hat{\rho}^2} = 0.27 \cdot \sqrt{1+0.82^2} = 0.35$. This value corresponds very closely with the results presented in Table 1 where the median of estimates $100\hat{\sigma}_e$ for individual economies is 0.31. Nonetheless, Table 1 also shows that there are considerable differences across individual countries as the estimated standard error $100\hat{\sigma}_e$ varies from 0.11 in Poland to 0.94 in Estonia.

Berger's (2011) estimate of volatility of the GDP cyclical component is $100\hat{\sigma}_e = 0.09$ (with a standard error of this estimate of 0.02). This is a slightly lower value than those presented in Table 1, which is probably caused by the fact that Berger worked with pre-crisis data.

Clark (1989) estimated the standard deviation of the random error in the equation for the output gap for the US economy as $100\hat{\sigma}_e = 0.73$. His estimate is more than two times higher than the median of the corresponding estimates for European countries presented in Table 1 and also than the value calculated above using the results reported by Chen and Mills (2012). Nonetheless, this does not mean that the variability of the output gap in the US economy is significantly higher than in Europe as the output gap in the USA is less persistent, as discussed above.

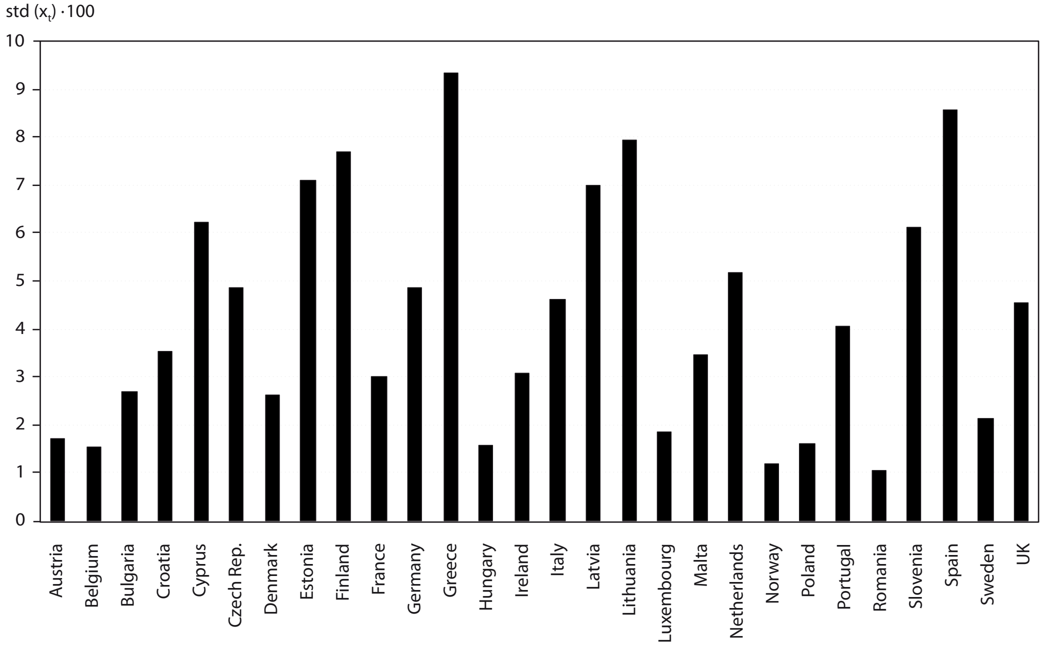
Output gap volatility might be measured by unconditional standard deviation of the AR(2) process (4), which is given by:

$$
std(x_t) = \sqrt{\frac{(1-\phi_2)\cdot \sigma_e^2}{(1+\phi_2)\cdot\left[(1-\phi_2)^2 - \phi_1^2\right]}}. \tag{11}
$$

Estimate of $std(x_t)$ is obtained by substituting the corresponding estimates $\hat{\phi}_1$, $\hat{\phi}_2$ and $\hat{\sigma}_e^2$ into the relation (11). Results are depicted in Figure 1. Unconditional standard deviation is multiplied by 100 in order to obtain more readable results.

The three countries with the highest output gap volatility are Greece, Spain and Lithuania. A surprising finding is that a substantially high variability of the gap is also detected for Finland. Low output gap volatility is observed in Austria, Belgium, Hungary, Luxembourg, Norway, Poland, Romania and Sweden. This finding is especially surprising for Romania, which has the lowest output gap volatility among all 28 studied economies. The case of Romania is substantially specific. As will be seen in subsequent paragraphs, Romania also has one of the highest volatilities of the growth rate g_t of the GDP trend component as measured by $100\sigma_w$ and also an extremely high volatility of the trend component n_t as measured by $100\sigma_v$. Movements in real GDP are thus mostly permanent as they influence primarily the trend components of real output.

Figure 1 Output gap volatility as measured by *std* (x_t) · 100



Source: Authors own calculations based on data obtained from Eurostat (2018)

Output gap volatility in the USA might be obtained by plugging Clark's (1989) estimates $\hat{\phi}_1 = 1.47$, $\hat{\phi}_2 = -0.59$ and $100\hat{\sigma}_e = 0.73$ into relation , which yields $std\left(x_t\right) \cdot 100 = 2.37$. The volatility of the cyclical component of output in the US economy is thus substantially low, even though $100\hat{\sigma}_e = 0.73$ is rather high. As already mentioned, this result is a direct consequence of low output gap persistency in the USA.

3.1.3 Volatility of the growth rate g_t of the GDP trend component (measured by $100\sigma_w$)

In contrast to the previous results obtained for parameter σ_e, the coefficient σ_w is statistically significant (at least at the 10% level of significance) only for approximately half of the studied countries – Bulgaria, Cyprus, the Czech Republic, France, Germany, Greece, Hungary, Latvia, Malta, Portugal, Romania, Spain and the United Kingdom. The growth rate of the trend component g_t is not constant over time in these countries as the standard deviation σ_w is significantly different from zero. Mostly, these are less developed European countries except for France, Germany and the United Kingdom, where $100\sigma_w$ is statistically different from zero, but attains considerably low values – France (0.04), Germany (0.09) and the United Kingdom (0.05). The highest values of $100\hat{\sigma}_w$ are detected for the Czech Republic (0.28), Hungary (0.22), Latvia (0.21) and Romania (0.25).

Literally a stable growth rate g_t of the GDP trend component (measured by $100\sigma_w$) is detected for Denmark (0.00), Ireland (0.00), Italy (0.00) and Poland (0.00). One might be tempted to interpret from these results that almost nothing will affect the growth rate of these economies in the long run – not even the economic crisis of 2008. Such an interpretation would of course be misleading as the long-run trend component of output is determined not only by $100\sigma_w$ but also by $100\sigma_v$ (to be discussed systematically later on). The most striking contrast between these two measures is observed for Ireland ($100\hat{\sigma}_w = 0$ is the lowest value among the studied economies and at the same time $100\hat{\sigma}_v = 2.82$ is the highest value among all 28 countries). A similar situation is also detected for Denmark and Poland.

Very low estimates of $100\hat{\sigma}_w$ are also obtained for Belgium (0.03), the Netherlands (0.03), Norway (0.03) and Sweden (0.02). These estimates correspond closely to the result reported by Kim and Nelson (1999) for the US economy ($100\hat{\sigma}_w = 0.03$, 0.02 being the standard error of this estimate) and by Chen and Mills (2012) for aggregated euro area data ($100\hat{\sigma}_w = 0.02$, with a standard error of 0.01). Their estimates are slightly lower than the results presented in this paper as the median of $100\hat{\sigma}_w$ for the 28 analysed European economies is 0.07.

3.1.4 Volatility of GDP trend component n_t (measured by $100\sigma_v$)

The coefficient σ_v measuring the volatility of the trend component is statistically significant at the 1% level in all countries except Belgium. The random error v_t influencing the GDP trend component n_t (without changing the growth rate g_t of this trend component) thus proved to have a highly significant effect in virtually all studied economies. Belgium is the only country where both standard deviations $100\hat{\sigma}_w = 0.03$ and $100\hat{\sigma}_v = 0.00$ describing the variability of the trend component of GDP are statistically insignificant and very low. The trend component of GDP is thus extremely stable in this country.

Berger (2011) estimated a similar UC model for aggregated European data from 1970 to 2005 and his estimate of $100\sigma_v$ is 0.45 (with a standard error of this estimate of 0.03). Kim and Nelson (1999) report their estimate for the US economy as $100\hat{\sigma}_v = 0.49$ (with a standard error of 0.06). These findings are slightly lower than the median 0.74 of estimates for the 28 individual countries presented in Table 1, which is probably caused by the fact that they did not analyse data during the huge economic crisis. The results in Table 1 also show that there are substantial differences in the volatility of the trend component of GDP (σ_v) across individual European countries.

The medians of the coefficients $100\hat{\sigma}_e$, $100\hat{\sigma}_w$ and $100\hat{\sigma}_v$ are 0.31, 0.07 and 0.74, respectively. A lower than median value for all three estimated parameters is observed for France, Italy, the Netherlands, Portugal and the United Kingdom. These countries show a stable development of the GDP trend n_t, and growth of the GDP trend g_t as well as the output gap x_t. The other extreme is economies with higher than median values for all three estimated parameters, which means that components n_t, g_t and x_t are rather volatile. This condition is satisfied for the following less developed countries: Estonia, Greece, Latvia, Lithuania, Malta and Slovenia.

3.2 Trend component of unemployment rate

The volatility of the trend component of the unemployment rate is measured by σ_ε. This parameter turned out to be significantly different from zero at least at the 5% level of significance for only about half of the studied economies: Austria, Belgium, Bulgaria, Cyprus, France, Hungary, Italy, Malta, Slovenia, Spain and the United Kingdom. This suggests that the variability of the unemployment rate trend is not as high as the volatility of the GDP trend component in quite a lot of European countries.

Berger (2011) estimated the coefficient $100\sigma_\varepsilon$ for the whole of Europe as 0.09 (with 0.01 being the standard error of this estimate). Clark (1989) reports the value for the US economy as $100\hat{\sigma}_\varepsilon = 0.17$. Empirical analysis performed in this paper shows that the estimate for a typical European countryis quite close to Berger's result as the median of estimates for 28 economies is equal to 0.10.

3.3 Okun's law

As far as the relation between output gap and cyclical unemployment rate is concerned, estimated parameters $\hat{\alpha}_0$ and $\hat{\alpha}_1$ mostly have a negative sign. In some (rather rare) cases, an estimate of $\hat{\alpha}_0$ or $\hat{\alpha}_1$ turned out to be positive. Nonetheless, for Okun's law to hold, it is sufficient that the sum $\hat{\alpha}_0 + \hat{\alpha}_1$ of the estimated coefficients is negative. This condition is satisfied in all 28 studied economies, thereby confirming the validity of Okun's law.

Statistical insignificance (at the 10% level) of both parameters α_0 and α_1 is observed for 10 economies: Bulgaria, Croatia, Cyprus, Hungary, Ireland, Luxembourg, Malta, Norway, Poland and Romania. The Czech Republic and Germany stand on the other side with both parameters α_0 and α_1 being statistically significant (at least at the 10% level of significance).

These results might be improved if the coefficient α_1 is a priori set to zero. Only the contemporary output gap x_t is then assumed to have an influence on the unemployment rate gap (C_t). This approach is taken, for example, by Ball et al. (2017). Statistical insignificance of the parameter α_0 (at the 10% level) was in this case detected only for Luxembourg, Malta, Norway and Romania. Statistical significance of α_0 at the 10% level was observed for two countries (Croatia and Poland). Okun's law parameter α_0 was significant at the 5% level in the cases of Belgium, Bulgaria, the Czech Republic, Denmark and Hungary. In the remaining 17 economies it turned out that the coefficient α_0 was statistically significant even at the 1% level of significance. This finding confirms the validity and strength of Okun's law across individual European economies. Cyclical properties of the unemployment rate are thus driven to a great extent by the output gap.

The output gap lagged one period (x_{t-1}) has a stronger negative influence on the unemployment rate than the current gap (x_t) in 20 of the monitored economies. Thus, only in eight cases is the reverse true (Greece, Ireland, Latvia, Lithuania, Luxembourg, Poland, Romania and Spain). This confirms that the unemployment rate is a lagging indicator (its current values depend mainly on the lagged output gap).

The estimated parameters of Okun's law are in line with other empirical studies for most countries. Kim and Nelson (1999) estimated Okun's law in a slightly modified form to specification (7) as they assumed that not only the current and one-period lagged output gap (x_t and x_{t-1}) but also the output gap lagged two periods x_{t-2} have an influence on the unemployment rate gap C_t. They report the following estimates (standard errors are in parentheses): $\hat{\alpha}_0 = -0.34\ (0.06)$, $\hat{\alpha}_1 = -0.16\ (0.03)$ and $\hat{\alpha}_2 = -0.07\ (0.01)$. The overall effect of the output gap on the gap of unemployment rate is thus $\hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 = -0.57$. Clark's (1989) estimates for the US economy are $\hat{\alpha}_0 = -0.33$ and $\hat{\alpha}_1 = -0.18$ with the overall effect $\hat{\alpha}_0 + \hat{\alpha}_1 = -0.51$. The median value for such an overall effect calculated from Table 1 is $median\left(\hat{\alpha}_0 + \hat{\alpha}_1\right) = -0.60$, which is quite close to the estimate obtained by Kim and Nelson (1999).

Mankiw (2012) posited that for the US economy a one per cent deviation of output from potential causes an opposite change in the unemployment rate of half a percentage point. This assertion is practically the same as Clark's (1989) estimate of $\hat{\alpha}_0 + \hat{\alpha}_1 = -0.51$ and also corresponds closely to the median value found by empirical investigation in this paper for 28 European economies.

Ball et al. (2017) estimated Okun's law (by ordinary least squares using annual data from 1980 to 2011) in the form:

$$U_t - U_t^* = \beta \cdot \left(y_t - y_t^*\right) + \varepsilon_t, \tag{12}$$

where U_t is the unemployment rate and y_t is (the log of) GDP, U_t^* and y_t^* are trend estimates obtained using the Hodrick-Prescott filter and ε_t is i.i.d. random error.

In the long run, the coefficient β is equivalent to $\alpha_0 + \alpha_1$ in the presented model (7). Furthermore, Clark's bivariate model (1)-(7) was also estimated here using the a priori assumption $\alpha_1 = 0$ in order to obtain results more comparable with regression (12). Ball et al. (2017) econometrically analysed relationship (12) for the US economy and for 20 advanced OECD countries. Comparison with estimates presented earlier in Table 1 (for countries analysed in this paper as well as by Ball et al., 2017) is summarized in the following Table 2. Standard errors of estimated coefficients are again shown in parentheses below estimated parameters, and the symbols ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Table 2 Comparison of results reported by Ball et al. (2017) ($\hat{\beta}$ in the first row) with findings of this paper ($\hat{\alpha}_0 + \hat{\alpha}_1$ and $\hat{\alpha}_0$ in the second and third row)

Country	Austria	Belgium	Denmark	Finland	France	Germany	Ireland	Italy
$\hat{\beta}$	−0.14*** (0.04)	−0.51*** (0.08)	−0.43*** (0.05)	−0.50*** (0.05)	−0.37*** (0.04)	−0.37*** (0.06)	−0.41*** (0.04)	−0.25*** (0.07)
$\hat{\alpha}_0 + \hat{\alpha}_1$	−0.29*** (0.07)	−0.27*** (0.09)	−0.51*** (0.15)	−0.59*** (0.09)	−0.46*** (0.06)	−0.43*** (0.09)	−1.35 (1.20)	−0.36*** (0.07)
$\hat{\alpha}_0 \;\; (\alpha_1 = 0)$	−0.26*** (0.09)	−0.19** (0.08)	−0.76** (0.31)	−0.69*** (0.14)	−0.55*** (0.14)	−0.62*** (0.15)	−0.51*** (0.18)	−0.40*** (0.09)

Country	Netherlands	Norway	Portugal	Spain	Sweden	UK
$\hat{\beta}$	−0.51*** (0.07)	−0.29*** (0.04)	−0.27*** (0.04)	−0.85*** (0.05)	−0.52*** (0.07)	−0.34*** (0.05)
$\hat{\alpha}_0 + \hat{\alpha}_1$	−0.52*** (0.09)	−0.81 (0.61)	−0.77*** (0.15)	−1.11*** (0.16)	−0.50*** (0.10)	−0.54*** (0.08)
$\hat{\alpha}_0 \;\; (\alpha_1 = 0)$	−0.59*** (0.13)	−1.13 (1.10)	−0.78*** (0.18)	−1.39*** (0.26)	−0.75*** (0.18)	−0.61*** (0.12)

Source: Authors own calculations based on data obtained from Eurostat (2018)

Firstly, Table 2 shows that Okun's law estimates $\hat{\alpha}_0 + \hat{\alpha}_1$ are slightly higher in absolute value than the corresponding values $\hat{\beta}$ reported by Ball et al. (2017). The only notable exception is Belgium (and Sweden with very similar results: $\hat{\beta} = -0.52$ and $\hat{\alpha}_0 + \hat{\alpha}_1 = -0.50$). Secondly, even slightly higher (in absolute value) estimates were obtained in most studied cases when the a priori setting was $\alpha_1 = 0$. Exceptions are Austria ($\hat{\alpha}_0 + \hat{\alpha}_1 = -0.29$, $\hat{\alpha}_0 = -0.26$), Belgium ($\hat{\alpha}_0 + \hat{\alpha}_1 = -0.27$, $\hat{\alpha}_0 = -0.19$) and Ireland ($\hat{\alpha}_0 + \hat{\alpha}_1 = -1.35$, $\hat{\alpha}_0 = -0.51$). Thirdly, the results obtained in this paper confirm the validity and strength of Okun's law reported by Ball et al. (2017) as Okun's law parameters are statistically significant even at the 1% level in most studied cases (exceptions are Norway and partly Ireland). Fourthly, standard errors of parameters estimated in this paper ($\alpha_0 + \alpha_1$ and α_0) are generally higher than the corresponding standard deviations of coefficients estimated by Ball (β). This is most clearly seen in the case of Norway and also Ireland. This is caused by the fact that Ball et al. (2017) treat the unobserved output gap and unemployment rate gap ($y_t - y_t^*$, $U_t - U_t^*$) as observable variables.

Ball et al. (2017) argue that reasonable values for β (for the US economy) should lie in the interval $(-1.5;0)$, which is satisfied in all studied cases. The highest (in absolute value) estimate of β is reported by Ball for Spain ($\hat{\beta} = -0.85$). Ball explains this strong negative influence of the output gap on the gap of unemployment rate by the prevalence of temporary employment contracts in Spain. The highest (in absolute value) estimate of $\alpha_0 + \alpha_1$ is detected for Ireland ($\hat{\alpha}_0 + \hat{\alpha}_1 = -1.35$). Nonetheless, this estimate is not statistically significant. The second highest value is observed for Spain ($\hat{\alpha}_0 + \hat{\alpha}_1 = -1.11$), confirming Ball's results. His finding relating to the strength of Okun's law in Spain is also confirmed by the highest (in absolute value) estimate of $\hat{\alpha}_0 = -1.39$.

Boďa et al. (2015) estimate Okun's law for Visegrad group countries by applying autoregressive distributed lag (ARDL) methodology using quarterly data from 1998 Q1 to 2014 Q2. Their long-run multiplier is comparable to the estimate of the overall effect $\hat{\alpha}_0 + \hat{\alpha}_1$ calculated here. Comparison of their findings with results presented in this paper is summarized in Table 3. Parameter standard deviations as well as the symbols *, ** and *** are indicated as before.

Table 3 Comparison of the long-run multiplier reported by Boďa et al. (2015) (the first row in the table) with findings of this paper ($\hat{\alpha}_0 + \hat{\alpha}_1$ in the second and third row)

Country	Czech Republic	Hungary	Poland
Long-run multiplier	−0.12*** (0.03)	−0.06 (0.04)	−0.08 (0.09)
$\hat{\alpha}_0 + \hat{\alpha}_1$	−0.54*** (0.14)	−1.49 (1.25)	−2.83 (1.86)

Source: Authors own calculations based on data obtained from Eurostat (2018)

Table 3 shows that findings presented in this paper imply a stronger negative relationship between the output gap and the gap of unemployment rate for these three countries than results reported by Boďa et al. (2015). Estimates of the long-run multiplier obtained by Boďa are unusually small (in absolute value). This is the case especially of Hungary and Poland, where estimates $\hat{\alpha}_0 + \hat{\alpha}_1$ calculated by the author of this paper are conversely substantially high. One robust finding reported here as well as by Boďa is that Okun's parameters are statistically insignificant in Hungary and Poland. Nonetheless, further empirical investigation of Okun's law for Visegrad group countries will be necessary.

Many empirical studies analyse Europe as a whole and not individual countries. Orlandi and Pichelmann (2000) estimated Okun's law for aggregated European annual data for 1960–1998 in the form:

$$\frac{\left(y_t - y_t^*\right)}{y_t^*} \cdot 100 = -\alpha \cdot \left(U_t - U_t^*\right) + \varepsilon_t, \tag{13}$$

where the variables have a similar interpretation to that in regression (12) except that GDP y_t is not in logs.

Parameter α from (13) relates to parameter β in according to $\beta \approx -1/\alpha$.[4] Orlandi and Pichelmann (2000) report the value $\hat{\alpha} = 1.8$ implying $\hat{\beta} = -1/1.8 = -0.56$, which is a value typically found for most European economies by Ball et al. (2017) as well as by empirical investigation performed in this paper.

Berger and Everaert (2008) formulated a UC model and applied Bayesian econometric techniques for aggregated European and US data for the period 1970 Q1–2003 Q4. Specifically, they estimated an equation for Okun's law in a form similar to (13) with the analogical parameter for α estimated for Europe as 2.15, which yields $\hat{\beta} = -1/2.15 = -0.47$.

Berger (2011) works with a similar UC model to the one applied in this paper with a Phillips curve added to the model specification. Berger's (2011) analysis is based on aggregated European data over the period 1970–2005. His equation for Okun's law is the same as specification (7) and reports the following estimation result: $\hat{\alpha}_0 = -0.49$ and $\hat{\alpha}_1 = -0.10$.

Novák and Darmo (2019) estimate Okun's law in a disaggregated manner for individual European economies by applying panel data econometric methods. Unfortunately, they assume that the parameter relating output with unemployment is the same for all countries. Moreover, they work with Okun's law in a differenced form that is not directly comparable with the specification applied in this paper. Specifically, Novák and Darmo (2019) estimate the following panel data regression:

$$\Delta U_{it} = \alpha_i + \beta \cdot gy_{it} + \gamma \cdot \Delta U_{i,t-1} + \varepsilon_{it}, \tag{14}$$

where ΔU_{it} is the first difference (a year-on-year change) of the unemployment rate in country i, gy_{it} represents real GDP growth in country i and ε_{it} is random error.

[4] The relation $\beta \approx -1/\alpha$ applies only approximately for two reasons. Firstly, there is a slightly different definition of the output gap in Formulas (12) and (13). Secondly, interchanging dependent and independent variables in the regression for Okun's law does not produce algebraically equivalent regression estimates (Plosser and Schwert, 1979).

Formula (12) considered by Ball et al. (2017) might be expressed in a differenced form if we assume constant trend components $U_t^* = U^*$ and $y_t^* = y^*$. Differencing (12) then yields:

$$U_t - U_{t-1} = \beta \cdot \left(y_t - y_{t-1} \right) + \varepsilon_t^* . \tag{15}$$

A comparison of relations (14) and (15) reveals that coefficient β relates GDP to unemployment rate in both cases. Thus, the coefficient β from (14) is in this sense comparable to the parameter β from (12). Novák and Darmo (2019) used data between 2001 and 2014 and report the value $\hat{\beta} = -0.29$, which is also in line with results presented and discussed earlier in this paper.

Novák and Darmo (2019) also found structural breaks in regression (14). Nonetheless, Brůha and Polanský (2015) argue that instability is often found when using a difference form of Okun's law and found evidence that the relationship is stable when cyclical components of output and unemployment are used instead of differences. Similarly, Ball et al. (2017) estimated Okun's law in its gap form and state that Okun's law is a strong and stable relationship in most countries, one that did not change substantially during the Great Recession. Ball also found that coefficient β in Okun's law (12) varies substantially across countries, which is confirmed by the results presented in this paper as well. Therefore, the assumption of constancy of this parameter across individual economies introduced by Novák and Darmo (2019) is not satisfied.

4 TREND-CYCLE DECOMPOSITON

Smoothed estimates of the GDP cycle (x_t), growth of the GDP trend component (g_t) and finally the trend component of the unemployment rate (L_t) together with the observed unemployment rate (U_t) are illustrated for all 28 European economies in Figure 2.

Figure 2 Smoothed estimates of output gap x_t, growth g_t of the GDP trend component and the trend component of the unemployment rate L_t together with the observed unemployment rate

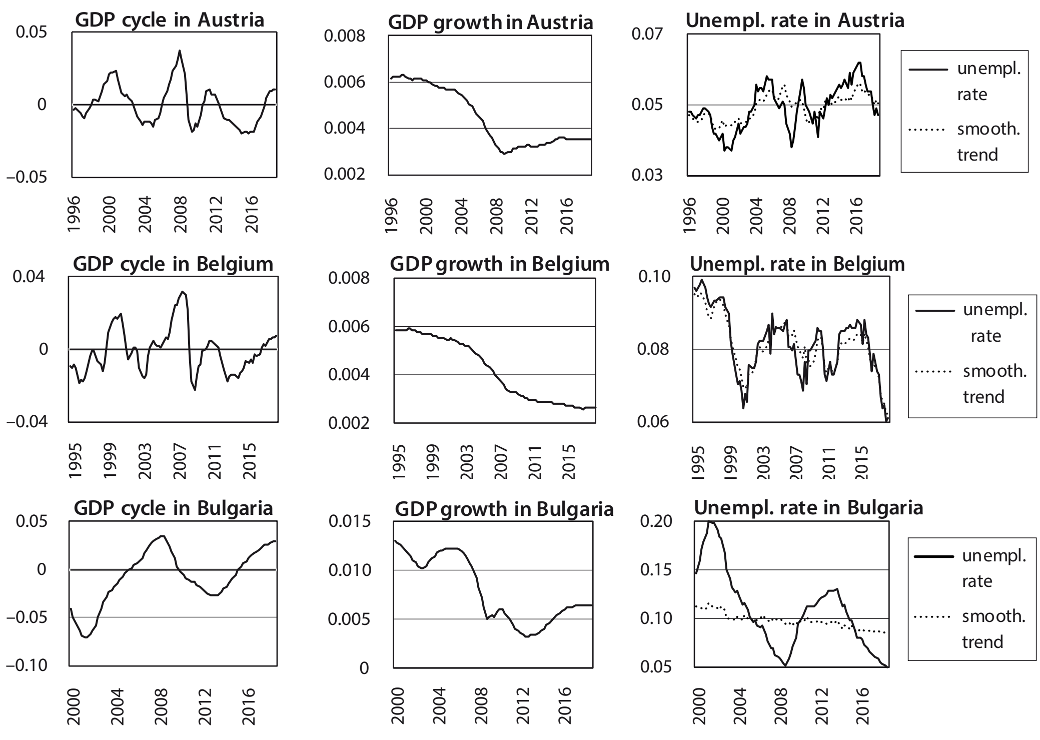


Figure 2 (continuation)

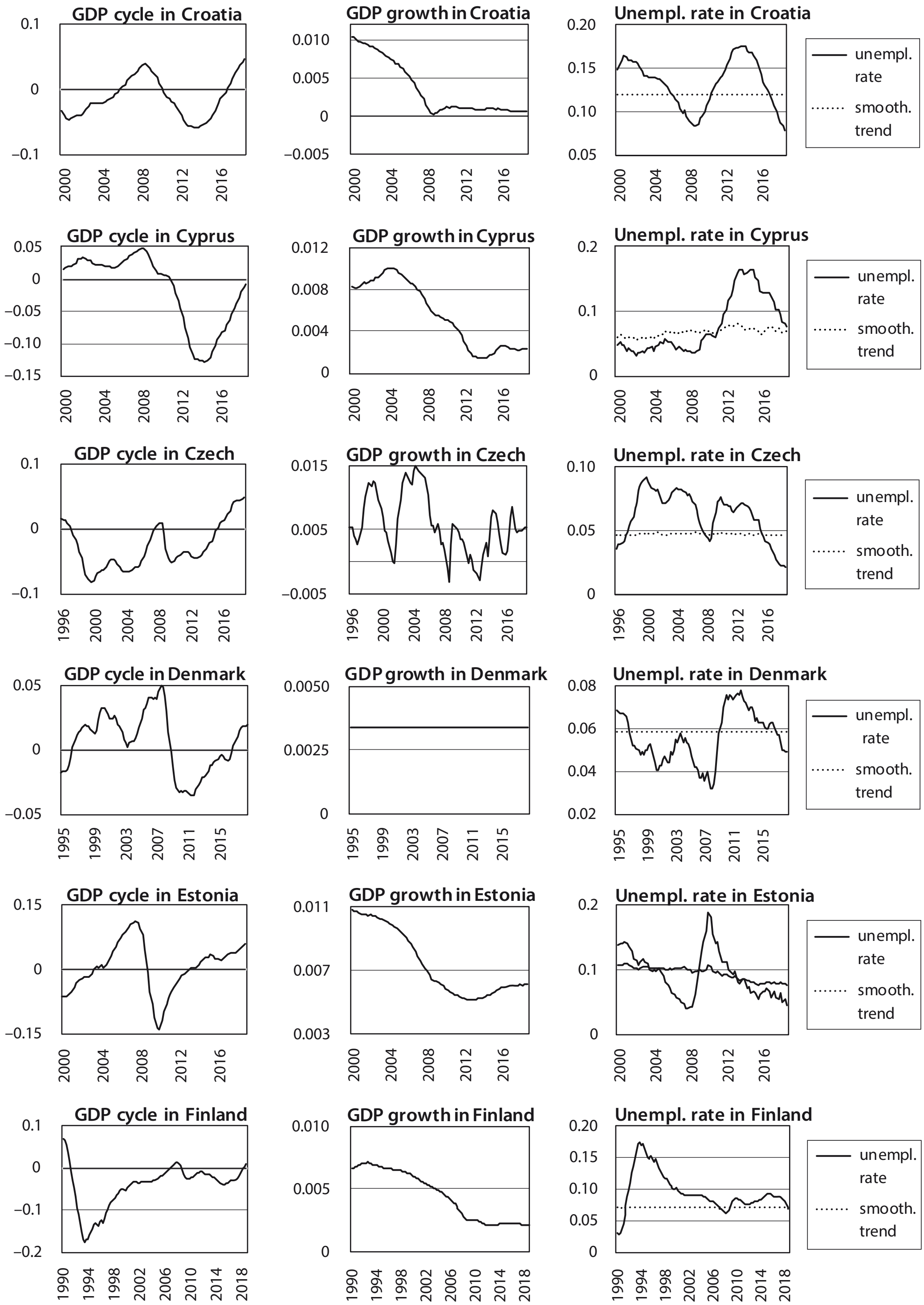


Figure 2 (continuation)

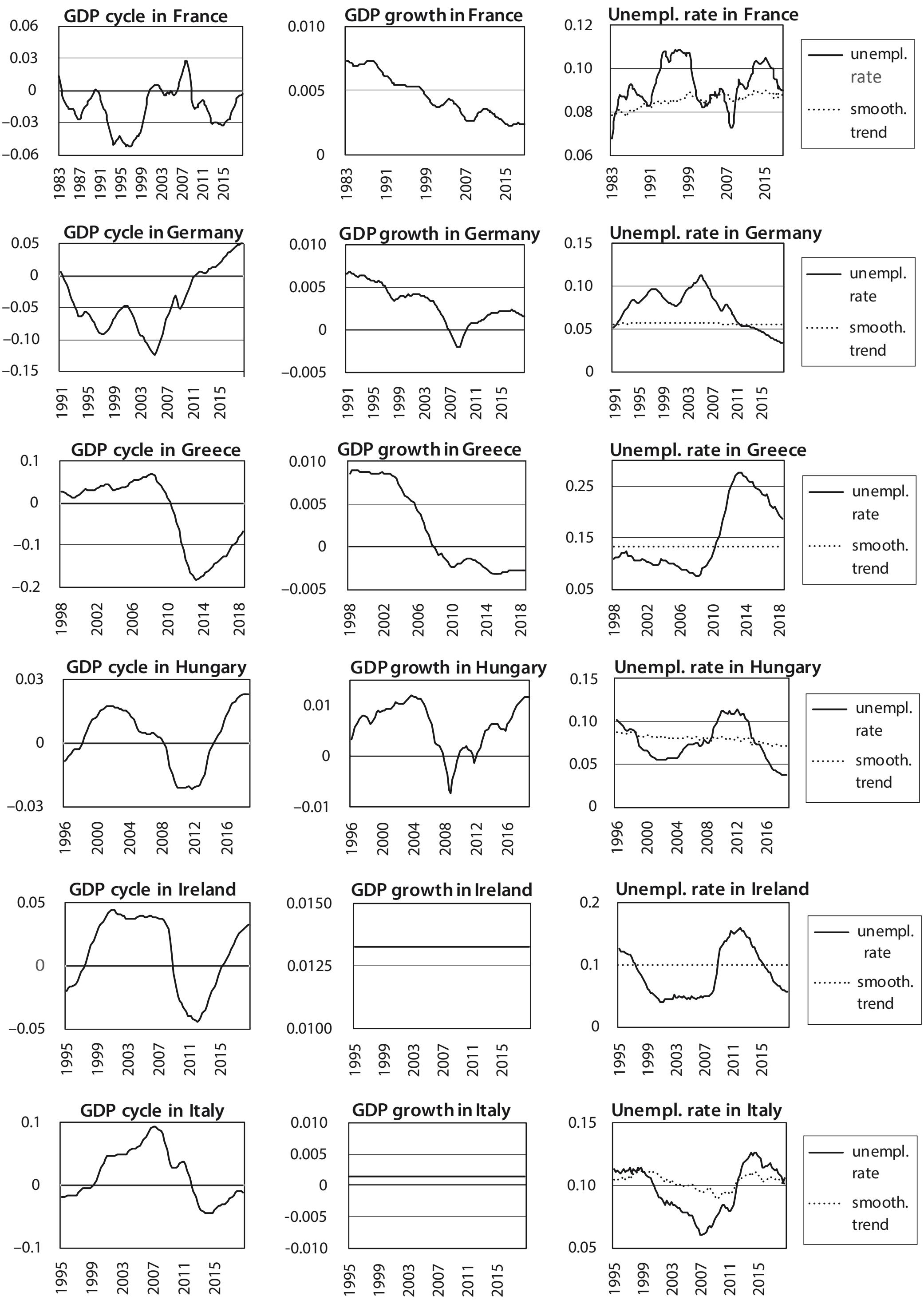


Figure 2 (continuation)

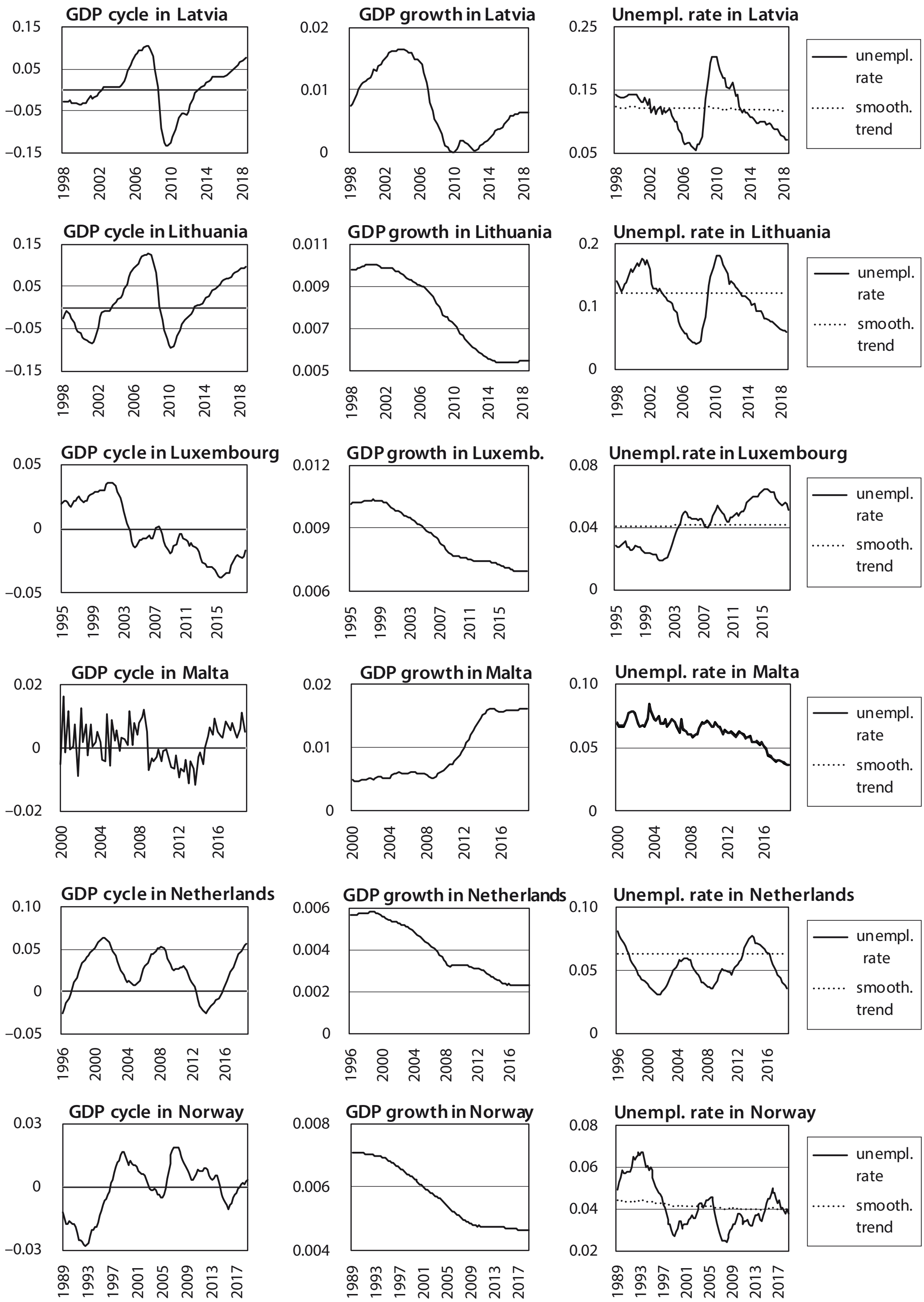


Figure 2 (continuation)

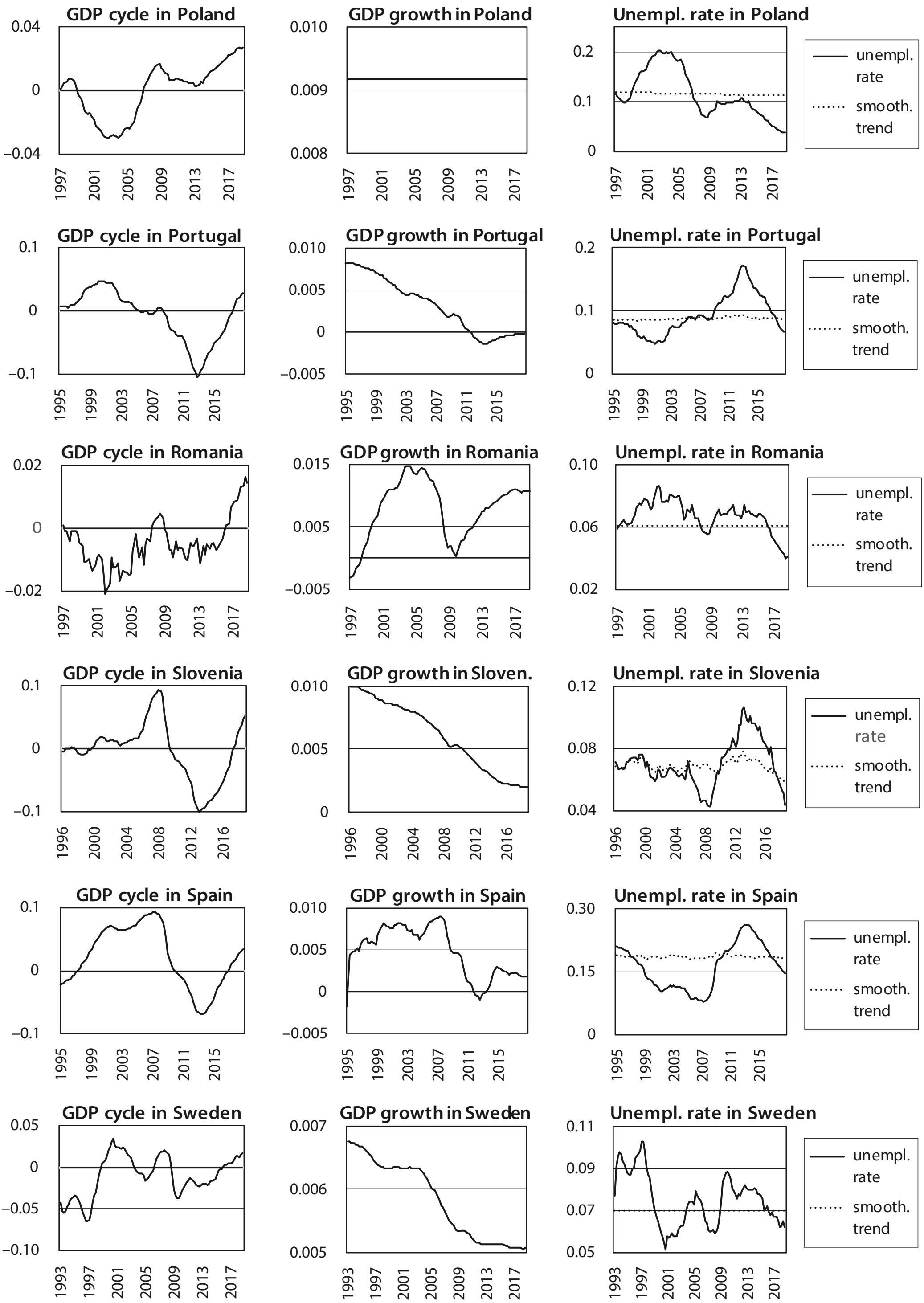


Figure 2 (continuation)

Source: Authors own calculations based on data obtained from Eurostat (2018)

4.1 GDP cycle

The cyclical component of GDP declined after the beginning of the crisis in 2008 in most countries. However, there are a few exceptions. The development of the gap in Finland seems untouched by the crisis, Germany experienced an increase in the output gap after 2008, and only a slight decrease has been observed for Malta, Norway, Poland, Romania and Sweden.

One important finding is that the output gap was very close to zero or even positive in 2018 for all countries except Greece and Luxembourg where GDP is approximately 7% and 2% below the trend value, respectively. Therefore, the huge long-lasting fall in GDP experienced by most countries is not a consequence of a decline in the output gap, but is caused by a decrease in the trend component of GDP.

A more detailed discussion on the output gap can be found in Section 5 where a comparison with HP-filtered estimates of the gap is described.

4.2 Growth of GDP trend component

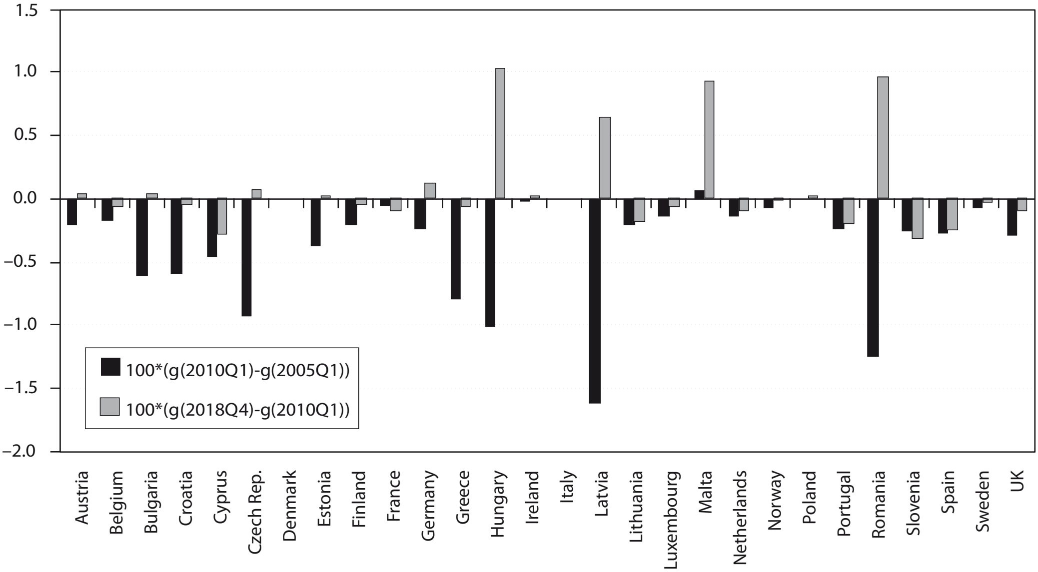
A decline in the growth g_t of the GDP trend component during the current economic crisis has been detected for most countries. Similar results were found by Lemoine et al. (2011), Ball (2014) and Haltmaier (2012). However, there are some exceptions. GDP growth in the Czech Republic was highly volatile with downturn as well as upturn movements. Malta even experienced an increase in the growth g_t of the GDP trend component. This makes the long-run development of GDP in this economy practically unaffected by the current economic crisis. Growth g_t was virtually constant during the studied periods in Denmark, Ireland, Italy and Poland.

These results correspond to the previous findings reported earlier in this paper in Section 3.1 when discussing estimates of $100\sigma_w$. The estimated volatility $100\hat{\sigma}_w$ of the random error w_t in Formula $g_t = g_{t-1} + w_t$ was found to be zero for these aforementioned countries (Denmark, Ireland, Italy and Poland). As already discussed above, zero estimates of $100\sigma_w$ in these countries (leading to the constant growth rate g_t observed in the second column of Figure 2) are compensated by high estimates of $100\sigma_v$, which applies especially for Ireland ($100\hat{\sigma}_v = 2.82$), but also for Poland ($100\hat{\sigma}_v = 0.95$) and Denmark ($100\hat{\sigma}_v = 0.76$).

Figure 3 summarizes the development of the growth g_t of the GDP trend component during the crisis and post-crisis periods in selected European economies. Specifically, a change (in percentage points) of g_t is calculated for the crisis period 2005 Q1–2010 Q1 and the post-crisis period 2010 Q1–2018 Q4 according to the relations $100 \cdot \left(g_{2010Q1} - g_{2005Q1} \right)$, $100 \cdot \left(g_{2018Q4} - g_{2010Q1} \right)$.

The most remarkable decline of g_t during the crisis period is detected for Latvia, Romania and Hungary. Fortunately, those countries hit most by the current crisis experienced an increase of g_t during the post-crisis period. A substantial decrease of g_t in the crisis period is also observed for Bulgaria, Croatia, the Czech Republic and Greece. Furthermore, Cyprus, Lithuania, Portugal, Slovenia and Spain suffered from a (moderate) decline of g_t in both crisis and post-crisis periods.

Figure 3 Decline in the growth g_t of the GDP trend component between 2005 Q1–2010 Q1 (crisis period) and 2010 Q1–2018 Q4 (post-crisis period)



Source: Authors own calculations based on data obtained from Eurostat (2018)

These results correspond closely to the findings previously discussed in this paper. Specifically, the growth g_t of the GDP trend component was found to be highly volatile (due to high and statistically significant σ_w) in the following countries: Bulgaria, Cyprus, the Czech Republic, Greece, Hungary, Latvia, Malta, Portugal, Romania and Spain. Comparison with Figure 3 shows that all these countries (except Malta) are economies facing problems with a decline of g_t.

Figure 2 shows that the growth g_t of the GDP trend component was even negative for some time in some countries, namely the Czech Republic, Germany, Greece, Hungary, Portugal and Spain. Nonetheless, it was in positive values in 2018 for all these countries except Greece where a negative growth rate g_t persisted until 2018Q4 (the last date in the data set). But it would be too soon to conclude that the long-run GDP growth g_t has already recovered. The variable g_t in 2018 attained much lower values (compared with the values of the pre-crisis period) for most of the economies.

4.3 Trend of unemployment rate

The smoothed unemployment rate trend is found to be constant for most economies. A slight variation of the trend L_t is observed for Austria, Belgium, Bulgaria, Estonia, France, Hungary, Italy, Malta, Slovenia and the United Kingdom. The case of Malta is specific in that the trend L_t closely corresponds to the observed unemployment rate U_t. A similar result is reported by Clark (1989) for Japan. An interesting finding is that some upward tendencies of L_t might be observed only in France. Downward tendencies of L_t are visible in Estonia, Hungary, Malta and the United Kingdom. These findings support the hypothesis that hysteresis effects have not played an important role.

The constant smoothed unemployment rate trend for most economies is an important and interesting result. It is surprising, especially for countries like Cyprus, Greece, Ireland, Latvia, Lithuania, Poland, Portugal and Spain, because the observed unemployment rate in these economies ranges from 0.05 to 0.25 (or even higher). Huge fluctuations of the observed unemployment rate U_t in these countries are thus caused by large swings in cyclical component C_t of the unemployment rate and not by movements in the trend L_t. This conclusion might seem at odds with economic intuition. The reason for this result is that the unemployment rate U_t has been decreasing systematically and significantly in the last few (5–8) years in these economies.

Similar results regarding the constancy of the trend component L_t were also obtained by Clark (1989) for West Germany and the United Kingdom. Clark reports that the estimated trend component L_t and observed unemployment rate U_t ranged between:

a) $L_t \in (0.045;0.075)$, $U_t \in (0.03;0.11)$ for the US economy (1947 Q1–1986 Q2),

b) $L_t \in (0.05;0.08)$, $U_t \in (0.03;0.13)$ for Canada (1955 Q1–1986 Q2),

c) $L_t = 0.015$, $U_t \in (0.01;0.09)$ for West Germany (1960 Q1–1986 Q2),

d) $L_t = 0.02$, $U_t \in (0.015;0.13)$ for the United Kingdom (1960 Q1–1986 Q2),

e) $L_t = U_t \in (0.01;0.03)$ for Japan (1960 Q1–1986 Q2).

According to Clark's (1989) findings, trend component L_t *L*turned out to be practically constant in West Germany and the United Kingdom despite the fact that the observed unemployment rate U_t spanned a considerably wide interval. Clark also reports that L_t and U_t practically coincide in Japan, which is a result that we detected for Malta. A surprising result is that the development of L_t and U_t, which would resemble Clark's (1989) findings for the USA and Canada (L_t is less volatile than U_t, but is not constant), was detected only for a few European economies. This suggests that individual European labour markets are fundamentally different to labour markets in the USA and Canada. One possible explanation is that the trend component L_t is influenced to a great extent by labour market institutions, which are not as flexible in Europe as in the USA or Canada.

5 COMPARISON WITH HP FILTER

The output gap estimated by Kalman smoother will now be compared with the gap obtained by applying an HP filter as a robustness check. Hodrick and Prescott (1997) proposed setting a smoothing parameter equal to 1 600 for quarterly data. This suggestion is based on their empirical investigation of the US economy. As individual European countries studied in this paper might be fundamentally different to the US economy, other values of the smoothing parameter suggested in literature (Baxter and King, 1999; Backus et al., 1992; Correia et al., 1992; Cooley and Ohanian, 1991) will be taken into account. This robustness check is important as Boďa et al. (2015) argue that an improperly chosen smoothing parameter might cause illusive cycles to appear. Other disadvantages of an HP filter are discussed by Boďa and Považanová (2019) and Plašil (2011).

For annual data, Baxter and King (1999) recommended a value of around 10, and Backus et al. (1992) advised setting a smoothing parameter equal to 100, whereas Correia et al. (1992), and Cooley and Ohanian (1991) proposed a value of 400. According to Ravn and Uhlig (1997), the smoothing parameter for quarterly data relates to its analogue for annual data according to the relation $\lambda_{quarter} = \lambda_{annual} \cdot 4^4$, which yields the following smoothing parameter values for quarterly data: 2 560, 25 600 and 102 400.

The output gap calculated by the HP filter with a smoothing parameter equal to 2 560 was virtually the same as for the most common value of 1 600. For this reason, only the values 1 600, 25 600 and 102 400 are taken into account. Figure 4 compares the output gap calculated by applying the HP filter with that obtained by Kalman smoother.

Figure 4 Estimates of output gap obtained by Kalman smoother and HP filter with smoothing parameter set equal to values of 1 600, 25 600 and 102 400

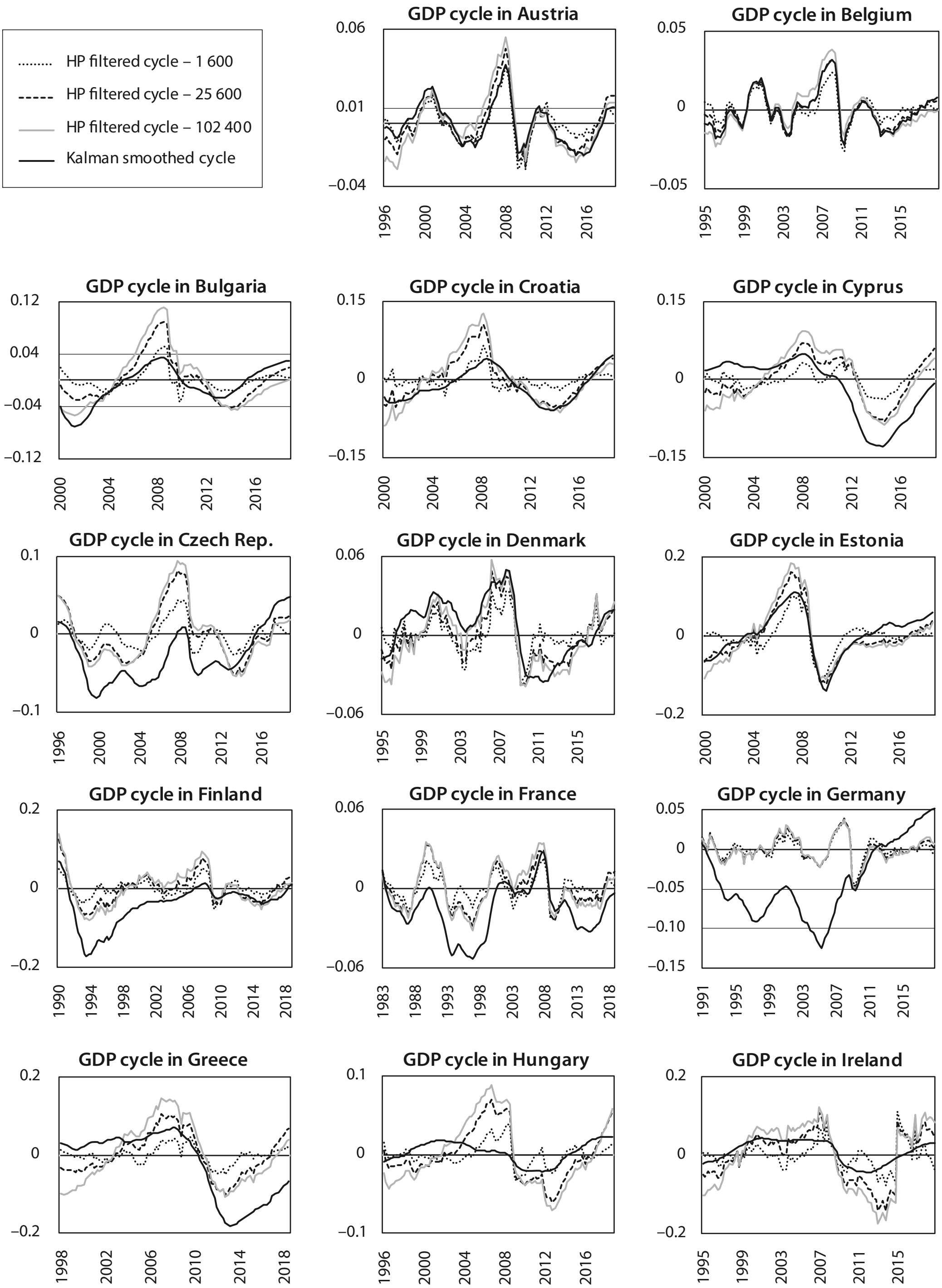
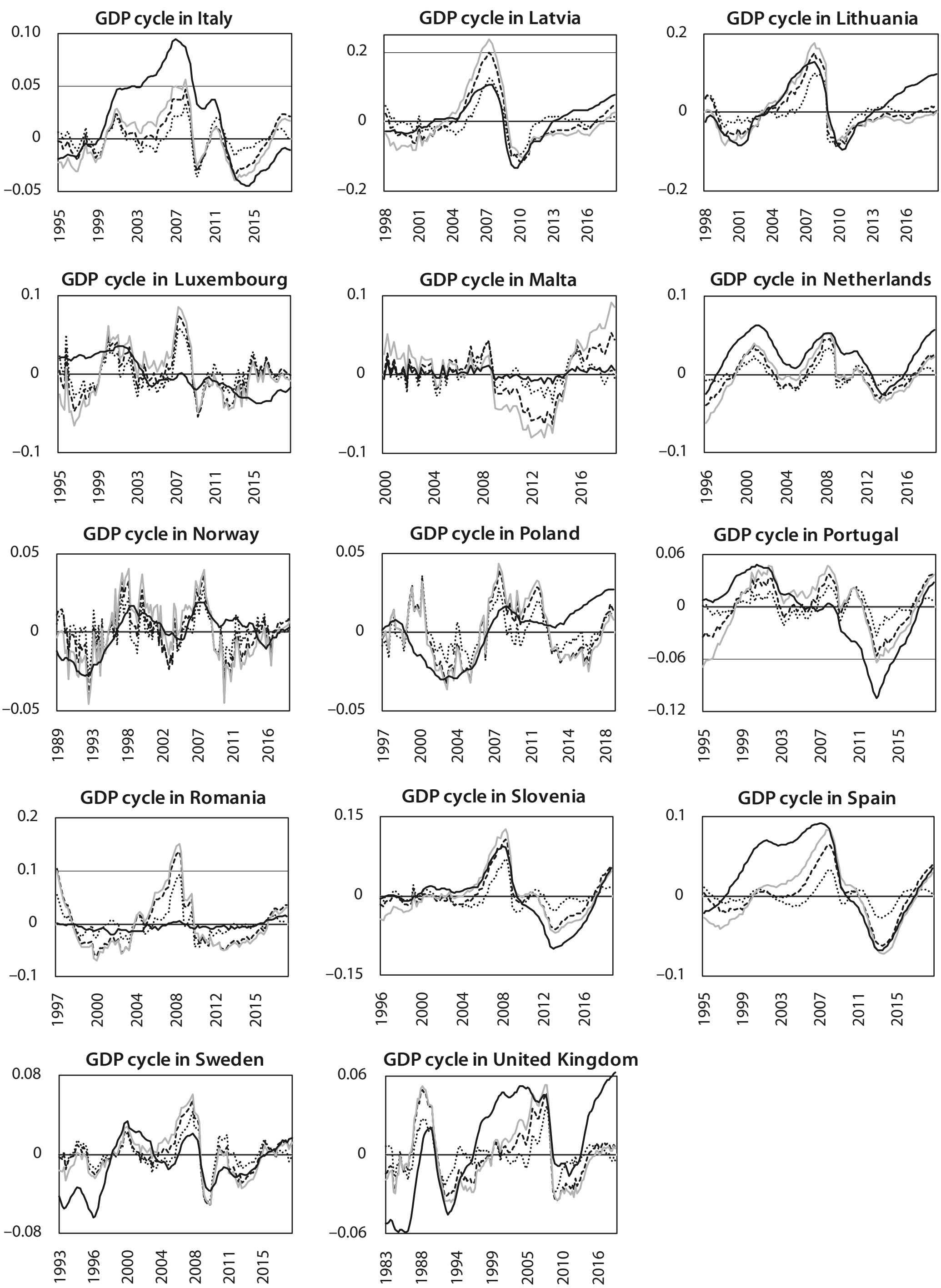


Figure 4 (continuation)



Source: Authors own calculations based on data obtained from Eurostat (2018)

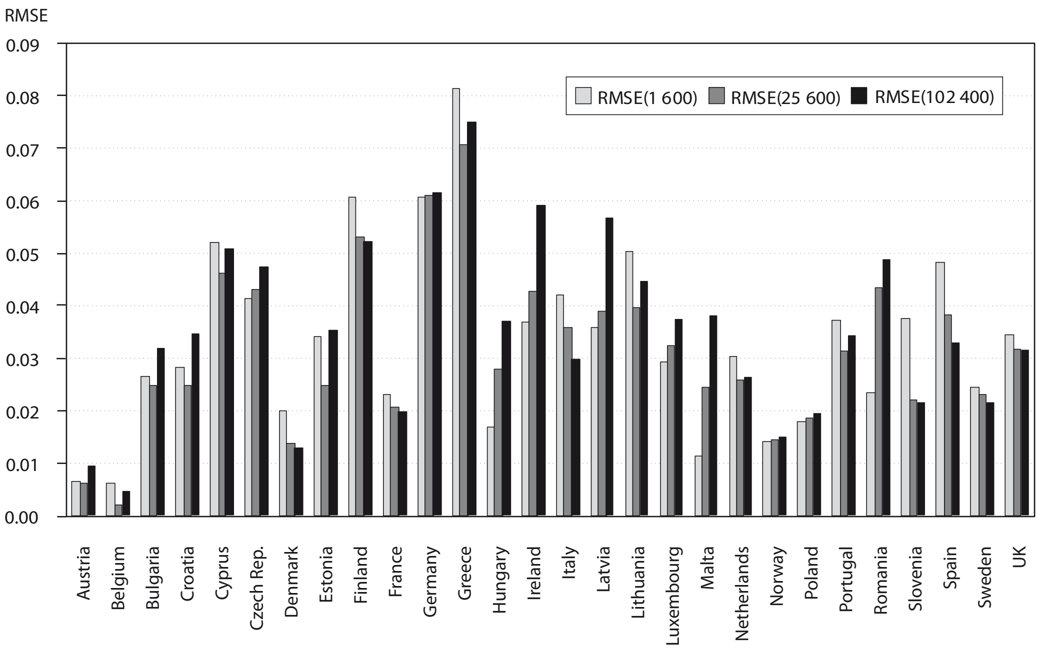
The (dis)similarity of the discussed output gap estimates can be measured by root mean squared error (RMSE) as follows:

$$RMSE(\lambda) = \sqrt{\frac{\sum_{t=1}^{n}\left(x_t^{Kalman} - x_t^{HP,\lambda}\right)^2}{n}},$$

where x_t^{Kalman} is the output gap obtained by Kalman smoother, $x_t^{HP,\lambda}$ represents the output gap calculated by HP filter with smoothing parameter λ, and n is the number of observations for a given economy.

RMSE was calculated for all 28 European economies with results summarized in Figure 5.

Figure 5 Comparison of output gap estimates x_t^{Kalman} and $x_t^{HP,\lambda}$ (λ ´= 1 600, 25 600, 102 400) using root mean squared error (RMSE)



Source: Authors own calculations based on data obtained from Eurostat (2018)

Figure 5 illustrates that different methods produced very similar results, especially for Austria, Belgium, Denmark, France, Norway, Poland and Sweden. The opposite is true mainly for Greece, where x_t^{Kalman} differs significantly from $x_t^{HP,\lambda}$ regardless of the smoothing parameter λ. Considerable differences between Kalman- and HP-filtered output gap measures are also detected for Cyprus, the Czech Republic, Finland, Germany, Ireland, Latvia, Lithuania and Romania.

Nonetheless, it can be seen from Figure 4 that a similar development of x_t^{Kalman} and $x_t^{HP,\lambda}$ is observed even in countries with the highest RMSE (Greece and Romania), i.e. time periods where the output gap is rising (decreasing) are mostly the same for all output gap estimation methods.

Moreover, there are some facts that proved to be robust across all the applied methods of estimation. Firstly, the output gap has already recovered since the crisis and is close to zero or even positive at the end of the sample in 2018 for most studied economies. Exceptions might be Greece, Italy and Luxembourg, where the HP filter suggests positive close-to-zero values of the output gap for all values of the smoothing parameter, but the Kalman smoother indicates a negative value of the gap. Secondly, the output gap experienced a decline after the beginning of the crisis in 2008 in all countries. There are a few exceptions to this when the gap is estimated by Kalman smoother (Germany, Malta, Poland and Romania). Nonetheless, HP-filtered estimates of the gap indicate a slump even in these countries.

Mostly, the Kalman smoother estimates the decrease in the gap after 2008 to be deeper and more prolonged than the HP filter. Nonetheless, this difference is smaller when the smoothing parameter is set to higher values.

The output gap calculated by HP filter with a smoothing parameter equal to 1 600 is usually very close to zero even in times of crisis, which seems unrealistic, especially for economies like Greece. The output gap in this country ranges only from –0.05 to 0.05, was already equal to zero in 2014 and GDP was already above its trend in 2018 according to this estimate. Such a situation is highly unrealistic given the huge problems in this economy. Setting higher values of the smoothing parameter leads to a much wider range within which the HP-filtered output gap in Greece oscillates. This makes the HP-filtered gap much closer to the gap estimated by Kalman smoother. A similar situation can also be seen in Cyprus, Denmark, Italy, Portugal, Slovenia and the United Kingdom.

Another unrealistic feature of the gap calculated by HP filter is that it was estimated to be positive at the beginning of the crisis in 2008 in all 28 studied economies. It is rather unrealistic that all 28 economies with mostly unsynchronized business cycle fluctuations would be in the same phase of the cycle at one specific moment. Output gaps obtained by the Kalman smoother have more realistic features as the (Kalman) smoothed gap had a negative value in Germany in 2008 and was roughly zero in Finland, Hungary, Luxembourg, Malta, Portugal and Romania.

CONCLUSION

An extensive empirical investigation has been carried out in this paper. Parameters of the well-known unobserved components Clark's (1989) model were estimated for 28 European economies. Consequently, the estimated model was applied in order to (1) empirically analyse cyclical and trend components of GDP and the unemployment rate, and (2) investigate the validity and strength of Okun's law.

Econometric estimation revealed that output has a highly persistent stationary cyclical component for all studied economies except Malta and Romania. Both cyclical and trend components in Malta are quite atypical and different to other European economies. Interestingly, there is a notable similarity between this economy and Japan when the results reported in this paper are compared with those obtained by Clark (1989). Comparison with Clark's (1989) estimates also confirms that the persistence of the gap in all European countries (apart from Malta and Romania) is much higher than in the US economy.

The negative impact of the current economic crisis on the long-run growth of the GDP trend turned out to be highest in Bulgaria, Croatia, the Czech Republic, Hungary, Greece, Latvia and Romania. Fortunately, those countries hit most by the current crisis (Hungary, Latvia and Romania) experienced an increase in the growth of the GDP trend component during the post-crisis period after 2010.

These results correspond closely to the findings regarding the variability of the growth of the GDP trend component – high and significant volatility was detected for Bulgaria, Cyprus, the Czech Republic, Greece, Hungary, Latvia, Malta, Portugal, Romania and Spain. The highest standard deviation value was observed for the Czech Republic, so the long-run growth in the Czech Republic is substantially variable.

These findings regarding substantial heterogeneity across individual countries emphasize the importance of analysing disaggregated European data. Such an approach taken in this paper represents a significant contribution to empirical literature as many studies use only aggregated European data (Azevedo et al., 2003; Berger and Everaert, 2008; Berger, 2011; Bernhofer et al., 2014; Chen and Mills, 2012; Galati et al., 2016; Lemoine et al., 2010; Orlandi and Pichelmann, 2000; Proietti, 2004).

Parameter estimation of the equation describing the negative relation between output gap and cyclical unemployment rate confirmed the validity of Okun's law as the total effect of the output gap on the gap of unemployment rate turned out to be negative in all studied economies. When assuming that only the contemporary output gap has an influence on unemployment, then (1) statistical insignificance of Okun's law coefficient at the 10% level was detected only in four economies (Luxembourg, Malta, Norway and Romania), and (2) statistical significance of Okun's law parameter at the 1% level was observed in 17 countries among all 28 studied economies. Cyclical properties of the unemployment rate are thus driven mainly by the cyclical component of GDP. We therefore confirm Ball et al.'s (2017) assertion that Okun's law is a strong and stable relationship in most countries, one that has not changed substantially during the current economic crisis.

Estimated coefficients in Okun's law turned out to be in line with other empirical studies. The median value for the overall effect of the output gap on the unemployment rate for the 28 analysed economies is –0.60, which is quite close to Clark's (1989) estimate of –0.51 as well as to Kim and Nelson's (1999) estimate of –0.57 for the US economy. Other empirical studies for aggregate euro area data usually report values within the interval (–0.6; –0.5) (Berger, 2011; Berger and Everaert, 2008; Orlandi and Pichelmann, 2000).

Another important finding is that estimated parameters in Okun's law vary quite substantially across countries, which is in line with Ball et al.'s (2017) results. Therefore, the assumption of constancy of this parameter across individual economies made in some empirical studies (e.g. Novák and Darmo, 2019) is not satisfied.

The volatility of the trend component of the unemployment rate turned out to be quite low. This was suggested by the fact that the standard deviation of the random error associated with this trend component was statistically significant in only approximately half of the studied economies. It was also supported by the calculated smoothed trend of the unemployment rate, which proved to be constant for most countries. Therefore, possible hysteresis effects in European labour markets have not played an important role yet. This is an interesting result, especially for economies like Cyprus, Estonia, Greece, Ireland, Latvia, Lithuania, Poland, Portugal and Spain. These countries exhibit huge swings in their unemployment rate, which are caused by movements in the cyclical component and not in trend. Similar findings were reported, for example, by Clark (1989) for West Germany and the United Kingdom.

The opposite is true for GDP because the huge long-lasting decline in GDP experienced by most countries is caused by a decline in the trend component of GDP and not in the cycle. The output gap decreased after 2008 in most countries (except Germany, Malta, Poland and Romania) but was very close to zero or positive in 2018 (except Greece). The long-lasting slump of GDP is thus caused by a decline in the trend.

A robustness check was performed for the output gap estimated by Kalman smoother because of the great economic importance of this variable. Specifically, the Hodrick-Prescott filter was applied and multiple values of the smoothing parameter were taken into account due to uncertainty about the correct value of this parameter for different European economies. The results obtained from the HP filter were quite similar to those obtained by the Kalman smoother, especially for Austria, Belgium, Denmark, France, Italy and Norway. The opposite is true mainly for Greece, which is an economy with non-standard economic conditions as this country is currently facing huge economic problems.

Some important facts proved to be robust across different methods – the output gap declined after the beginning of the crisis in 2008 in virtually all countries, but the gap has already recovered and was close to zero or even positive at the end of the sample in 2018. The output gap obtained from the HP filter also had some unrealistic features. This makes the estimates obtained by the Kalman smoother preferable, which supports the view given by Harvey and Jaeger (1993) regarding the superiority of the unobserved components approach to simple detrending methods. In some cases, unrealistic features of the HP-filtered gap were overcome by increasing the smoothing parameter above the most commonly used value of 1 600 (for quarterly data), which demonstrates that this commonly applied value might not be appropriate for all European economies.

The formulated model could be extended mainly in two ways. Firstly, the Phillips curve describing the relation between the output gap and inflation could be added into the model as done by Kloudová (2013), Orlandi and Pichelmann (2000), Beneš and N'Diaye (2004), Berger and Everaert (2008) and Berger (2011). Secondly, structural breaks in some variables as well as correlations between shocks to the unemployment rate and corresponding shocks to output might be considered as in Berger (2011), Chen and Mills (2012) and Proietti (2004).

ACKNOWLEDGEMENT

This work was processed with the contribution of the long-term institutional support of research activities by the Faculty of Informatics and Statistics, University of Economics, Prague (IP400040) and was supported by the Internal Grant Agency of the University of Economics, Prague, under Grant F4/53/2019.

References

AZEVEDO, J., KOOPMAN, S., RUA, A. *Tracking Growth and the Business Cycle: A Stochastic Common Cycle Model for the Euro Area*. Tinbergen Institute Discussion Paper, No. TI 2003-069/4, 2003.

BACKUS, D. K., KEHOE, P. J., KYDLAND, F. E. International Real Business Cycles. *Journal of Political Economy*, 1992, 100(4), pp. 745–775.

BAXTER, M. AND KING, R. Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series. *The Review of Economics and Statistics*, 1999, 81(4), pp. 573–593.

BALL, L. Long-Term Damage from the Great Recession in OECD Countries. *European Journal of Economics and Economic Policies: Intervention*, 2014, 11(2), pp. 149–160.

BALL, L., LEIGH, D., LOUNGANI, P. Okun's Law: Fit at 50? *Journal of Money, Credit and Banking*, 2017, 49(7), pp. 1413–1441.

BARRO, R. J. *Economic Growth in East Asia Before and After the Financial Crisis*. NBER Working Paper, No. W8330, Cambridge, Massachusetts: National Bureau of Economic Research, 2001.

BENEŠ, J. AND N'DIAYE, P. *A Multivariate Filter for Measuring Potential Output and the NAIRU: Application to the Czech Republic*. IMF Working Paper, No. 04/45, International Monetary Fund, 2004.

BERGER, T. AND EVERAERT, G. Unemployment Persistence and the NAIRU: A Bayesian Approach. *Scottish Journal of Political Economy*, 2008, 55(3), pp. 281–299.

BERGER, T. Estimating Europe's Natural Rates. *Empirical Economics*, 2011, 40(2), pp. 521–536.

BERNARDI, M. AND DI RUGGIERO, A. *Extracting the Italian Output Gap: a Bayesian Approach.* arXiv:1411.4898 [Stat.AP], Cornell University Library, 2014.

BERNHOFER, D., AMADOR, O, GÄCHTER, M., SINDERMANN, F. Finance, Potential Output and the Business Cycle: Empirical Evidence from Selected Advanced and CESEE Economies. *Focus on European Economic Integration*, 2014, 2, pp. 52–75.

BOĎA, M., MEDVEĎOVÁ, P., POVAŽANOVÁ, M. (A)symmetry in Okun's Law in the Visegrad Group Countries (in Slovak). *Politická ekonomie*, 2015, 63(6), pp. 741–758.

BOĎA, M. AND POVAŽANOVÁ, M. Okun's Law in the Visegrád Group Countries. *Europe-Asia Studies,* 2019, 71(4), pp. 608–647.

BRŮHA, J. AND POLANSKÝ, J. *Empirical Analysis of Labor Markets over Business Cycles: An International Comparison*. Working Papers 2015/15, Czech National Bank, Research Department, 2015.

CERRA, V. AND SAXENA, S. C. Did Output Recover from the Asian Crisis? *IMF Staff Papers*, 2005, 52(1), pp. 1–23.

CERRA, V. AND SAXENA, S. C. Growth Dynamics: The Myth of Economic Recovery. *American Economic Review*, 2008, 98(1), pp. 439–457.

CHEN, X. AND MILLS, T. C. Measuring the Euro Area Output Gap Using a Multivariate Unobserved Components Model Containing Phase Shifts. *Empirical Economics*, 2012, 43(2), pp. 671–692.

CLARK, P. K. The Cyclical Component of U.S. Economic Activity. *The Quarterly Journal of Economics*, 1987, 102, pp. 197–814.

CLARK, P. K. Trend Reversion in Real Output and Unemployment. *Journal of Econometrics*, 1989, 40, pp. 15–32.

COOLEY, T. J. AND OHANIAN, L. E. The Cyclical Behavior of Prices. *Journal of Monetary Economics*, 1991, 28(1), pp. 25–60.

CORREIA, I. H., NEVES, J. L., REBELO, S. T. Business Cycles from 1850–1950: New Facts about Old Data. *European Economic Review*, 1992, 36(2–3), pp. 459–467.

DUPASQUIER, C., GUAY, A., ST-AMANT, P. A Survey of Alternative Methodologies for Estimating Potential Output and the Output Gap. *Journal of Macroeconomics*, 1999, 21(3), pp. 577–595.

EUROSTAT. *Database–Eurostat* [online.]. [cit. 2.6.2019] <http://ec.europa.eu/eurostat>.

GALATI, G., HINDRAYANTO, I., KOOPMAN, S., VLEKKE, M. *Measuring Financial Cycles with a Model-Based Filter: Empirical Evidence for the United States and the Euro Area*. DNB Working Paper, No. 495, 2016.

HÁJEK, M. AND BEZDĚK, M. *Estimation of Potential Product and Output Gap in the Czech Republic* (in Czech). ČNB Working Paper, No. 26, 2000.

HÁJKOVÁ, D. AND HURNÍK, J. Cobb-Douglas Production Function: The Case of a Converting Economy. *Czech Journal of Economics and Finance*, 2007, 57(9–10), pp. 465–476.

HALTMAIER, J. *Do Recessions Affect Potential Output?* International Finance Discussion Paper 1066, Federal Reserve Board, 2012.

HARVEY, A. C. *Forecasting, Structural Time Series Models and the Kalman Filter*. 1st Ed. Cambridge: Cambridge University Press, 1989.

HARVEY, A. C. AND JAEGER, A. Detrending, Stylized Facts and the Business Cycle. *Journal of Applied Econometrics*, 1993, 8(3), pp. 231–247.

HODRICK, R. J. AND PRESCOTT, E. C. Postwar U.S. business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking,* 1997, 29(1), pp. 1–16.

JEMEC, N. *Output Gap in Slovenia. What Can We Learn from Different Methods?* Banka Slovenije Working Paper, No. 4/2012, 2012.

KIM, C. AND NELSON, C. R. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. Cambridge: The MIT Press, 1999.

KLOUDOVÁ, D. Output Gap as Indicator of Inflation – Case for Czech Economy (in Czech). *Politická ekonomie*, 2013, 61(55), pp. 639–652.

LEMOINE, M., DE LA SERVE, M. E., CHETOUANE, M. *Impact of the Crisis on Potential Growth: An Approach Based on Unobserved Components Models* (in French). Banque de France Working Paper, No. 331, 2011.

LEMOINE, M., MAZZI, G. L., MONPERRUS-VERONI, P. A. New Production Function Estimate of the Euro Area Output Gap. *Journal of Forecasting*, 2010, 29(1–2), pp. 29–53.

MANKIW, G. *Principles of macroeconomics.* 6th Ed. Mason, OH: South-Western Cengage Learning, 2012.

NOVÁK, M. AND DARMO, L. Okun's Law over the Business Cycle: Does It Change in the EU Countries after the Financial Crisis? *Prague Economic Papers*, 2019, 28(2), pp. 235–254.

OCHOTNICKÝ, P. Production Function Choice by Potential Output Estimation (in Slovak). *Ekonomický časopis/Journal of Economics*, 2008, 56(8), pp. 800–815.

OKUN, A. M. *Potential GNP: Its Measurement and Significance*. Reprinted as Cowles Foundation Paper 190, 1962.

ORLANDI, F. AND PICHELMANN, K. *Disentangling Trend and Cycle in the EUR-11 Unemployment Series: An Unobserved Component Modelling Approach*. Brussels: European Commision, 2000.

PLAŠIL, M. Potential Product, Output Gap and Uncertainty Rate Associated with Their Determination while Using the Hodrick-Prescott Filter (in Czech). *Politická ekonomie*, 2011, 59(4), pp. 490–507.

PLOSSER, C. AND SCHWERT, W. Potential GNP: Its Measurement and Significance: A Dissenting Opinion. In: BRUNNER, K. AND MELTZER, A. eds. *Three Aspects of Policymaking, Knowledge, Data and Institutions*, Amsterdam: Carnegie-Rochester Conference Series on Public Policy, North-Holland, 1979, pp. 179–186.

PROIETTI, T. State Space Decomposition under the Hypothesis of Non-Zero Correlation between Trend and Cycle with an Application to the Euro-zone. In: MAZZI, L. L. AND SAVIO, H. eds. *Monographs of Official Statistics*, Papers and Proceedings of the Third Colloquium on Modern Tools for Business Cycle Analysis, Research in Official Statistics, Eurostat, Luxembourg, 2004, pp. 292–325.

RAVN, M. AND UHLIG, H. On Adjusting the HP-Filter for the Frequency of Observations. *The Review of Economics and Statistics*, 2002, 84(2), pp. 371–380.

VOLOS, C. AND HADJIXENOPHONTOS, A. *An Unobserved Components Model Approach to the Relationship between Real GDP and Unemployment for Cyprus*. Neapolis University of Paphos Working Paper, No. 5, 2015.

ZIMKOVÁ, E. AND BAROCHOVSKÝ, J. Estimation of Potential Product and Output Gap in Slovak Conditions (in Czech). *Politická ekonomie*, 2007, 55(4), pp. 473–489.

Estimation of the Optimal Parameter of Delay in Young and Lowe Indices in the Fisher Index Approximation

Adam Juszczak [1] | *University of Lodz, Lodz, Poland*

Abstract

The Cost of Living Index (COLI) enables to show changes in the cost of household consumption assuming the constant utility level. The most commonly used way to approximate COLI is the Consumer Price Index (CPI) calculated by using the Laspeyres index. Many economists consider superlative indices such as the Fisher index as the best proxy for the COLI. However, it uses quantity data not only from a base but also the current period, which limits its usefulness. Thus, the indices like the Lowe index and the Young Index are used in order to approximate the Fisher index value without using current period expenditure data. Both of these indices use an additional parameter of delay. The purpose of this paper is to examine the influence of the parameter mentioned above on the Fisher index approximation using the empirical and simulation data.

Keywords

CPI, Young index, Lowe index, Laspeyres index, Fisher index, COLI, Cost of Living Index, Consumer Price Index, inflation

JEL code

C43, C49

INTRODUCTION

As an approximation of changes in the costs of household consumption assuming the constant utility (Cost of Living Index known as COLI), the Consumer Price Index is the most common way to measure inflation. The Cost of Living Index for a single household can be defined as the minimum cost of achieving a certain standard of living during a given period, divided by the minimum cost of achieving the same standard of living during a base period. However, in practice, the CPI is measured by the Laspeyres index, which is a subject of wide criticism. It risks bias due to ignoring changes in consumers' behavior (such as changing the retailers to these with lower prices) due to the price change, which results in overstating inflation. Thus, some economists treat the Laspeyres index as the Cost of Goods index (in opposite

[1] University of Lodz, Department of Statistical Methods, Fakulty of Economics, 90-255 Łódź, 3/5 POW Street, Poland. E-mail: adam.juszczak2@gmail.com.

to the Cost of Living Index). According to "superlative indices" theory developed by W. Erwin Diewert some indices such as the Fisher index can provide a fair approximation of the COLI "using the quantities in the base period as well as in the current reference period as weights in a symmetric fashion". Unfortunately, the Fisher index requires quantity data set from the current period, which takes time to process. This causes the inability of using the Fisher index results in many economic decisions such as monetary policy or adjusting social pensions. On the grounds of this issue statisticians proposed indices that approximate the Fisher index without using current expenditure data i.e. the AG mean index, the Lloyd-Mounton Index or the Lowe and the Young indices. The Lowe and Young indices compare two points in time, let us say 0 (base period) and τ, which can be any point between 0 and current period t, as well as precedes 0. The purpose of this paper is to approximate the optimal estimation of the τ parameter and verify the quality of obtained approximations. To reach this aim we realize empirical and simulation studies.

The structure of the paper is as follows: Section 1 discusses the connection between the Cost of Living Index and the Fisher index. Sections 2 and 3 introduce the Lowe price index and the Young price index. Section 4 describes some other approximations of the Fisher price index. Section 5 presents the simulation study, which concerns the bias of the previously mentioned indices. Section 6 displays an empirical study for 7 European countries and the EU benchmark for the 2006–2018 period. Last section demonstrates the main conclusions.

1 ROLE OF THE FISHER PRICE INDEX THE COLI MEASUREMENT

The COLI was introduced in 1961 by a committee chaired by George Stigler, which highlighted the difference between the CPI, in a form that was used then, and the true cost of living. The committee concluded by recommendation to the National Bureau of Labor Statistics in the USA to start using the COLI and adapt the Consumer Price index to obtain a better approximation of the Cost of Living index. Thirty-five years later in 1996, the Booskin Committee assessed the measurement of the COLI by the CPI in the US and concluded that it was overstating the true COLI value by 1.1 percent annually.

To define the Cost of Living index let us consider household preferences over commodities being represented by the utility function U(q) which is dual to the consumer expenditure function $E(P, \bar{u}) = \min_Q\{P^TQ|U(Q) \geq \bar{u}\}$. Most of households, wants to maximize the utility function for given budget limitations (in other words to minimize expenditure needed to achieve the utility level ū), and it leads to the following form of the Konüs price index:

$$P_K = \frac{E\left(P^T, \bar{u}\right)}{E\left(P^S, \bar{u}\right)}, \tag{1}$$

where *s* denotes the base period, *t* denotes for the current period and *P* considers prices at any moment *τ* are given by $P^T = [P_1^\tau, P_2^\tau, \ldots, P_N^\tau]^T$

The difference between the Cost of Living index which captures the changes of commodities quantity and the Laspeyres index that relies on quantities from the previous period is called the substitution bias and it has the biggest factor in miscalculating inflation rate. It is worth mentioning that even though in theory the Cost of Living index was defined by Russian economist Konüs in 1924, in practice the Fisher index is considered the easiest way to calculate COLI (Fisher, 1922).

As it was stated in the introduction, as a rule the Laspreyres index overstates true inflation because its formula takes under consideration quantities only from the previous time period:

$$P_{La} = \frac{\sum_{i=1}^{n} p_{i,t} q_{i,b}}{\sum_{i=1}^{n} p_{i,b} q_{i,b}}, \quad (2)$$

where $p_{i,t}$ means the price of a commodity i at current time moment t, $p_{i,b}$ – the price of commodity i at base time moment 0 and $q_{i,b}$ – the quantity of a commodity i at base time moment 0.

On the other hand, the Pasche index understates inflation because it takes only the quantity from the current period i.e.

$$P_{Pa} = \frac{\sum_{i=1}^{n} p_{i,t} q_{i,t}}{\sum_{i=1}^{n} p_{i,b} q_{i,t}}, \quad (3)$$

where $q_{i,t}$ means the quantity of commodity i at the current time moment t.

Because the Laspeyres index and the Pasche index have contrary biases, the Fisher index can be calculated as a geometric mean of them, i.e.

$$P_F = \sqrt{P_{La} P_{Pa}}. \quad (4)$$

2 LOWE PRICE INDEX

As it was mentioned above, the biggest flaw of the current price indices is the time needed for their publication. This time gap necessary to gather and process data causes low usefulness in economic decisions. That is why we use proxies for the Fisher Index.

Let us introduce some new period τ which precedes base period (b) (some authors (Białek, 2017) consider also situations when τ > b). The Lowe price index can be expressed as follows:

$$P_{LO} = \frac{\sum_{i=1}^{N} p_i^t q_i^\tau}{\sum_{i=1}^{N} p_i^b q_i^\tau} = \sum_{i=1}^{N} w_i^{\tau,b} \left(\frac{p_i^t}{p_i^b}\right), \quad (5)$$

where:

$$w_i^{\tau,b} = \frac{p_i^b q_i^\tau}{\sum_{k=1}^{N} p_k^b q_k^\tau}. \quad (6)$$

The arithmetic form of the Lowe index is not the only one. There is also a geometric version of this price index, i.e.

$$P_{GLO} = \prod_{i=1}^{N} \left(\frac{p_i^t}{p_i^b}\right)^{w_i^{\tau,b}}. \quad (7)$$

3 YOUNG PRICE INDEX

The second considered proxy for the Fisher price index is the Young index. The Young index is considered weaker in terms of fulfilled axioms, however, in some cases, it gives better Fisher index approximation than the Lowe Index (Armknecht and Silver, 2012). The Young index can be written as follows:

$$w_0^\tau = \frac{p_i^\tau q_i^\tau}{\sum_{k=1}^{N} p_k^\tau q_k^\tau}, \quad (8)$$

where:

$$w_0^\tau = \frac{p_i^\tau q_i^\tau}{\sum_{k=1}^{N} p_k^\tau q_k^\tau}. \qquad (9)$$

Similarly to the Lowe index case, we also take into consideration the geometric version of the Young index, i.e.

$$P_{GY} = \prod_{i=1}^{N} (\frac{p_i^t}{p_i^b})^{w_i^\tau}. \qquad (10)$$

4 OTHER PROXIES FOR THE FISHER INDEX FORMULA

The indices described in Sections 2 and 3 are not only those that can be used to approximate the Fisher index. We should also mention about the Arithmetic-Geometric (AG) mean index and the Lloyd-Moulton Index.

The AG mean index was proposed by Alan H. Dorfman and Janice Lent (2009), hence from their last names, it is sometimes called the L-D index as well. In the base version, it is the weighted from arithmetic mean of the Laspeyre s index and it's geometric counterpart i.e.

$$P_{LD} = \sigma \prod_{i=1} (\frac{p_i^t}{p_i^b})^{s_i^b} + (1-\sigma) \sum_{i=1} (\frac{p_i^t}{p_i^b}) s_i^b, \qquad (11)$$

where a parameter σ is elasticity of substitution of commodities covered, s_i^b is the expenditure share at base time 0 of the i-th commodity, i.e.

$$s_i^b = \frac{p_i^b q_i^b}{\sum_{i=1}^{N} p_i^b q_i^b} \quad . \qquad (12)$$

The second index which should be referred to is the Lloyd–Moulton (Lloyd, 1975; Moulton, 1996) index:

$$P_{LM} = \left\{ \sum_{i=1}^{N} s_i^b \left(\frac{p_i^t}{p_i^b} \right)^{1-\sigma} \right\}^{\frac{1}{1-\sigma}}, \qquad (13)$$

where parameter has the identical meaning as before (see Formula (11)).

The Lloyd – Moulton index has also an alternative version which was suggested by Huang, Waruna and Polard (2015) i.e.

$$P_{ModLM} = \left\{ \sum_{i=1}^{N} s_i^\tau \left(\frac{p_i^t}{p_i^b} \right)^{1-\sigma} \right\}^{\frac{1}{1-\sigma}}, \qquad (14)$$

where:

$$s_i^\tau = \frac{p_i^\tau q_i^\tau}{\sum_{i=1}^{N} p_i^\tau q_i^\tau}. \qquad (15)$$

5 SIMULATION STUDY

Through the simulation, we wish to check how the bias between the Fisher index and the studied indices differ for various delay parameters and product baskets. We consider several case studies, which differ from each other with respect to correlation between prices and quantities, the direction of price changes and inflation rate.

Case 1

Let us consider a scenario with N = 10 commodities where both prices and quantities change linearly in the following way:

$$p_i^T = p_i^b + \left(p_i^t - p_i^b\right)T\ , \tag{16}$$

$$q_i^T = q_i^b + \left(q_i^t - q_i^b\right)T,\ \mathrm{T} \in [0,1], \tag{17}$$

where p_i^b is goods price in the base period 0, p_i^t is the price in the current period t, q_i^b is the goods quantity in the base period and q_i^t is the quantity in current period. In this scenario, we are going to control the parameter of delay (τ) and we tend to optimize its value.

We selected four baskets for the simulation:

a) N = 10 goods with negative correlation between prices and quantities (prices increase and quantities decrease).

Table 1 The values of prices and quantities at time *0* and *t* for the case a

Goods no.	p^0	p^t	q^0	q^t
1	100	120	1 000	950
2	10	11	9 000	8 000
3	5	6.6	12 500	12 000
4	1 000	1 200	2 02	150
5	120	150	2 500	2 000
6	500	550	2 000	1 900
7	150	155	2 000	1 900
8	1 550	2 000	100	70
9	2 000	2 200	200	150
10	7	10	1 450	1 000

Source: Own construction

b) N = 10 goods negative correlation between prices and quantities (prices decrease and quantities increase).

Table 2 The values of prices and quantities at time *0* and *t* for the case b

Goods no.	p^0	p^t	q^0	q^t
1	100	95	1 000	1 100
2	10	9	9 000	9500
3	5	4.6	12 500	13 000
4	1 300	1 200	202	240
5	120	110	2 500	3 200
6	500	470	2 000	2 300
7	150	145	2 000	2 100
8	1 550	1 400	100	120
9	2 000	1 900	200	230
10	7	5	1 450	1 600

Source: Own construction

c) N = 10 with mixed goods. In five cases the price increased and quantity decreased and in three cases the price decreased and quantities increased. Hence, these can be considered normal goods. For one of the commodity, the price decrease was followed with a quantity decrease as well (which can be observed in some kinds of commodities such as computer games or gaming consoles, when the majority of purchases are made right after the introduction of the commodity to the market), and for the last one, the price increase caused the quantity increase (which is common for luxury goods – see Veblen paradox, 1899). As both of these cases are in minority in the consumer price index, basket there are represented as a minority in the simulation as well.

Table 3 The values of prices and quantities at time *0* and *t* for the case c

Goods no.	p^0	p^t	q^0	q^t
1	100	105	1 000	900
2	10	11	9 000	8 500
3	5	5.8	12 500	11 000
4	1 300	1 370	202	170
5	120	130	2 500	2 300
6	500	470	2 000	2 200
7	150	140	2 000	2 250
8	1 550	1 400	100	130
9	2 000	2 300	200	230
10	7	5	1 450	1 300

Source: Own construction

d) N = 10 with prices increase and quantity decrease with the aim for inflation around 2.5% (optimal parameter of inflation rate for the National Polish Bank).

Table 4 The values of prices and quantities at time *0* and *t* for the case d

Goods no.	p^0	p^t	q^0	q^t
1	100	102	1 000	970
2	10	10.4	9000	8 000
3	5	5.6	12 500	12 000
4	1 000	1 030	202	200
5	120	122	2 500	2 400
6	500	510	2 000	1 960
7	150	153	2 000	1 990
8	1 550	1 600	100	90
9	2 000	2 050	200	170
10	7	7.5	1 450	1 440

Source: Own construction

In the simulation, we changed the value of τ in the range [–2; 0.75]. Even though the most common practice is to use τ that precedes the base period, in some cases in previous studies τ parameter that was between the base and current period gave the best results.

Case 2

Using the same goods and services basket as in case one, let us consider exponential price and quantity change, i.e.

$$p_i^T = p_i^b \left(\frac{p_i^t}{p_i^b} \right)^T, \tag{18}$$

$$q_i^T = q_i^b \left(\frac{q_i^t}{q_i^b} \right)^T, \text{ where: } T \in [0,1]. \tag{19}$$

5.1 Simulation Results

Case 1a

Table 5 The values of the considered price indices for the case 1a

τ parameter	Laspeyres	Paasche	Fisher	Young	Geo.Young	Lowe	Geo.Lowe
–2.00	1.1389	1.1341	1.1365	1.1264	1.1242	1.1454	1.1425
–1.00	1.1389	1.1341	1.1365	1.1348	1.1322	1.1425	1.1397
–0.75	1.1389	1.1341	1.1365	1.1361	1.1335	1.1417	1.1389
–0.50	1.1389	1.1341	1.1365	1.1373	1.1346	1.1408	1.1380
–0.25	1.1389	1.1341	1.1365	1.1382	1.1355	1.1399	1.1371
0.25	1.1389	1.1341	1.1365	1.1393	1.1366	1.1378	1.1351
0.50	1.1389	1.1341	1.1365	1.1396	1.1369	1.1367	1.1340
0.75	1.1389	1.1341	1.1365	1.1397	1.1369	1.1354	1.1328

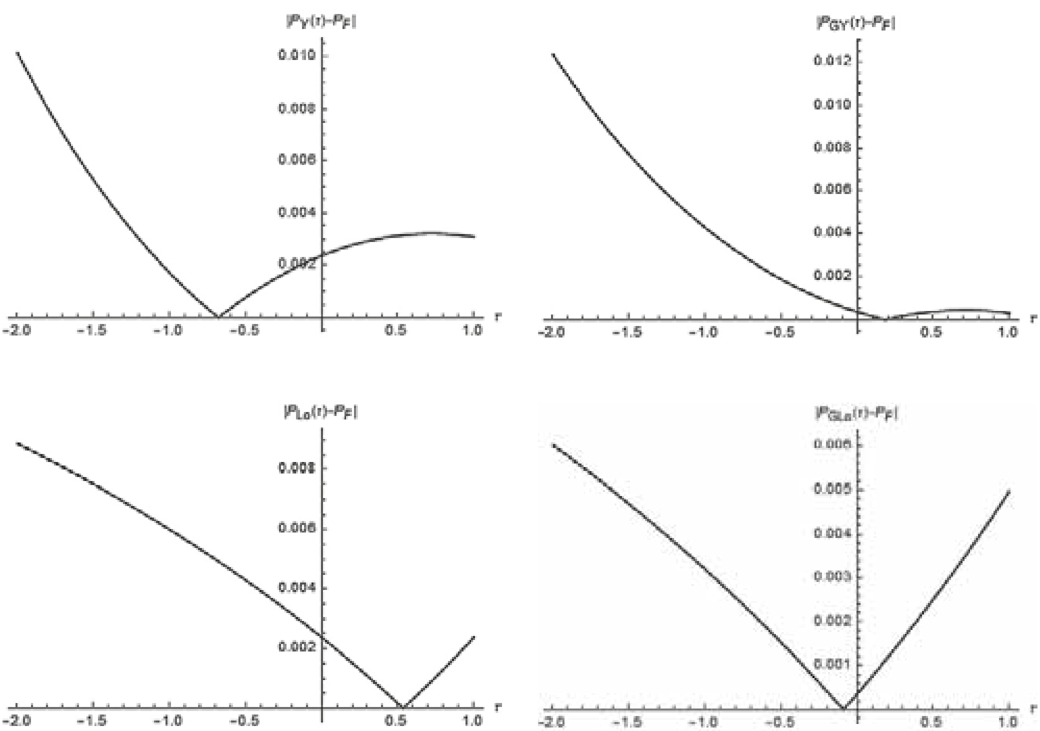
Source: Own construction in Mathematica 11

Table 6 Distance between considered price indices and the Fisher index for the case 1a

τ parameter	P_L-P_F	P_Y-P_F	P_{GY}-P_F	P_{Lo}-P_F	P_{GLo}-P_F
–2.00	0.0023675	–0.0101307	–0.0123396	0.0088750	0.0060260
–1.00	0.0023675	–0.0017211	–0.0042852	0.0059991	0.0031995
–0.75	0.0023675	–0.0003513	–0.0029695	0.0051726	0.0023886
–0.50	0.0023675	0.0007769	–0.0018858	0.0042952	0.0015283
–0.25	0.0023675	0.0016793	–0.0010194	0.0033620	0.0006139
0.25	0.0023675	0.0028495	0.0001008	0.0013054	–0.0013984
0.50	0.0023675	0.0031295	0.0003657	0.0001687	–0.0025092
0.75	0.0023675	0.0032086	0.0004360	–0.0010508	–0.0036996

Source: Own construction in Mathematica 11

Figure 1 Absolute differences between the Fisher index and the considered price indices as functions of *τ* (case 1a)



Source: Own construction in Mathematica 11

Case 1b

Table 7 The values of the considered price indices and their distances to the Fisher price index for the case 1b

τ parameter	Laspeyres	Paasche	Fisher	Young	Geo.Young	Lowe	Geo.Lowe
–2.00	0.9358	0.9353	0.9355	0.9365	0.9362	0.9375	0.9372
–1.00	0.9358	0.9353	0.9355	0.9360	0.9357	0.9365	0.9362
–0.75	0.9358	0.9353	0.9355	0.9359	0.9356	0.9363	0.9360
–0.50	0.9358	0.9353	0.9355	0.9359	0.9356	0.9361	0.9358
–0.25	0.9358	0.9353	0.9355	0.9358	0.9355	0.9359	0.9357
0.25	0.9358	0.9353	0.9355	0.9358	0.9355	0.9356	0.9354
0.50	0.9358	0.9353	0.9355	0.9358	0.9355	0.9355	0.9352
0.75	0.9358	0.9353	0.9355	0.9358	0.9355	0.9354	0.9351

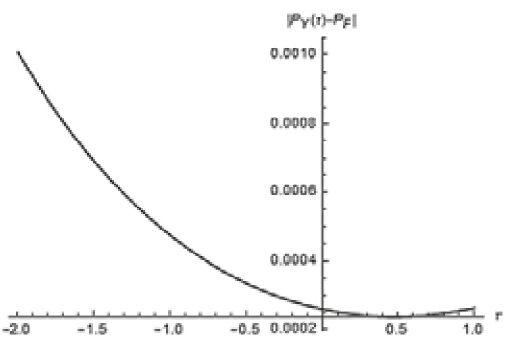
Source: Own construction in Mathematica 11

Table 8 Distance between considered price indices and the Fisher index for the case 1b

τ parameter	P_L-P_F	P_Y-P_F	P_{GY}-P_F	P_{Lo}-P_F	P_{GLo}-P_F
–2.00	0.0002600	0.0010074	0.0006373	0.0019821	0.0016669
–1.00	0.0002600	0.0004750	0.0001485	0.0009664	0.0006666
–0.75	0.0002600	0.0003967	0.0000801	0.0007671	0.0004703
–0.50	0.0002600	0.0003358	0.0000287	0.0005841	0.0002903
–0.25	0.0002600	0.0002906	–0.0000071	0.0004156	0.0001245
0.25	0.0002600	0.0002427	–0.0000369	0.0001157	–0.0001705
0.50	0.0002600	0.0002379	–0.0000327	–0.0000183	–0.0003023
0.75	0.0002600	0.0002449	0.0000168	–0.0001432	–0.0004251

Source: Own construction in Mathematica 11

Figure 2 Absolute differences between the Fisher index and the considered price indices as functions of τ (case 1b)



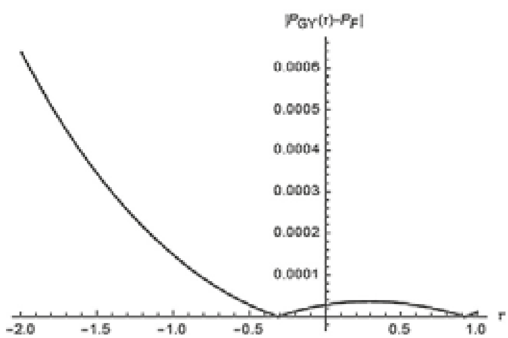
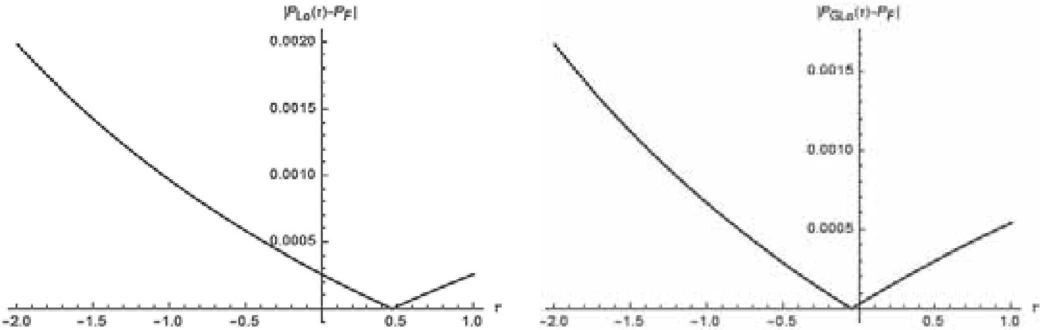


Figure 2 (continuation)



Source: Own construction in Mathematica 11

Case 1c

Table 9 The values of the considered price indices and their distances to the Fisher price index for the case 1c

τ parameter	Laspeyres	Paasche	Fisher	Young	Geo.Young	Lowe	Geo.Lowe
−2.00	1.00942	1.0053	1.0074	1.0038	1.0002	1.0192	1.0155
−1.00	1.00942	1.0053	1.0074	1.0061	1.0024	1.0140	1.0102
−0.75	1.00942	1.0053	1.0074	1.0069	1.0031	1.0128	1.0090
−0.50	1.00942	1.0053	1.0074	1.0077	1.0038	1.0117	1.0078
−0.25	1.00942	1.0053	1.0074	1.0085	1.0046	1.0105	1.0066
0.25	1.00942	1.0053	1.0074	1.0104	1.0064	1.0084	1.0044
0.50	1.00942	1.0053	1.0074	1.0114	1.0073	1.0073	1.0033
0.75	1.00942	1.0053	1.0074	1.0124	1.0083	1.0063	1.0023

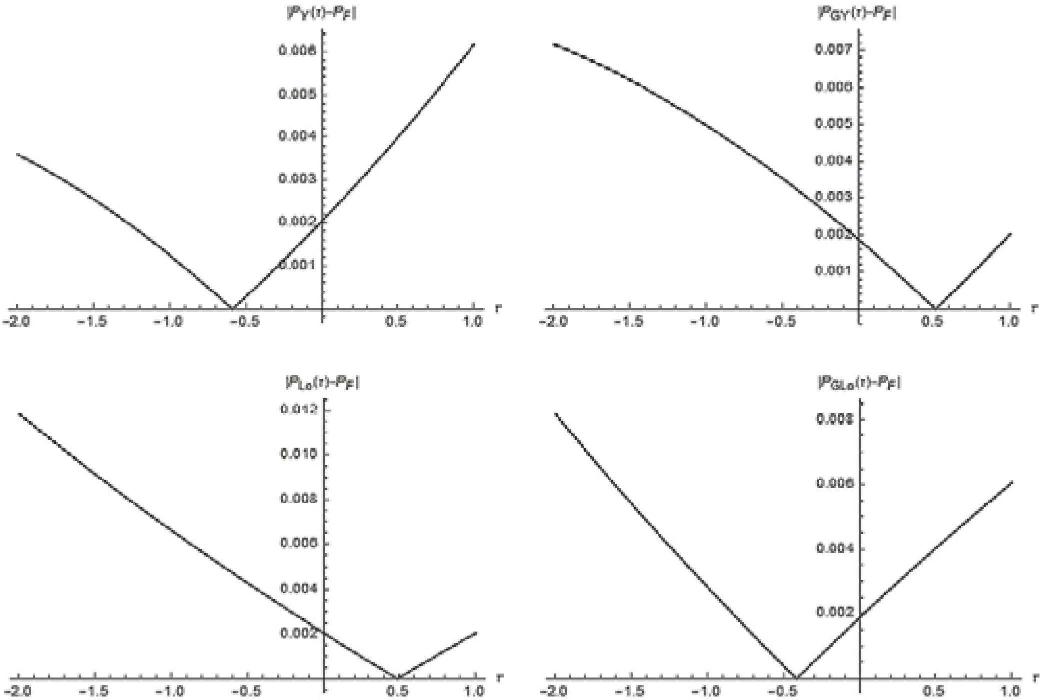
Source: Own construction in Mathematica 11

Table 10 Distance between considered price indices and the Fisher index for the case 1c

τ parameter	P_L-P_F	P_Y-P_F	P_{GY}-P_F	P_{Lo}-P_F	P_{GLo}-P_F
−2.00	0.0020476	−0.0036002	−0.0071668	0.0118328	0.0081712
−1.00	0.0020476	−0.0012475	−0.0049855	0.0066402	0.0028221
−0.75	0.0020476	−0.0005061	−0.0042910	0.0054400	0.0015898
−0.50	0.0020476	0.0002914	−0.0035417	0.0042756	0.0003956
−0.25	0.0020476	0.0011432	−0.0027394	0.0031452	−0.0007623
0.25	0.0020476	0.0030031	−0.0009812	0.0009811	−0.0029754
0.50	0.0020476	0.0040087	−0.0000276	−0.0000554	−0.0040336
0.75	0.0020476	0.0050635	0.0009744	−0.0010631	−0.0050615

Source: Own construction in Mathematica 11

Figure 3 Absolute differences between the Fisher index and the considered price indices as functions of τ (case 1c)



Source: Own construction in Mathematica 11

Case 1d

Table 11 The values of the considered price indices and their distances to the Fisher price index for the case 1d

τ parameter	Laspeyres	Paasche	Fisher	Young	Geo.Young	Lowe	Geo.Lowe
–2.00	1.02515	1.0250	1.0251	1.0248	1.0247	1.0253	1.0252
–1.00	1.02515	1.0250	1.0251	1.0250	1.0249	1.0253	1.0251
–0.75	1.02515	1.0250	1.0251	1.0250	1.0249	1.0252	1.0251
–0.50	1.02515	1.0250	1.0251	1.0251	1.0250	1.0252	1.0251
–0.25	1.02515	1.0250	1.0251	1.0251	1.0250	1.0252	1.0251
0.25	1.02515	1.0250	1.0251	1.0252	1.0251	1.0251	1.0250
0.50	1.02515	1.0250	1.0251	1.0252	1.0251	1.0251	1.0250
0.75	1.02515	1.0250	1.0251	1.0253	1.0251	1.0251	1.0250

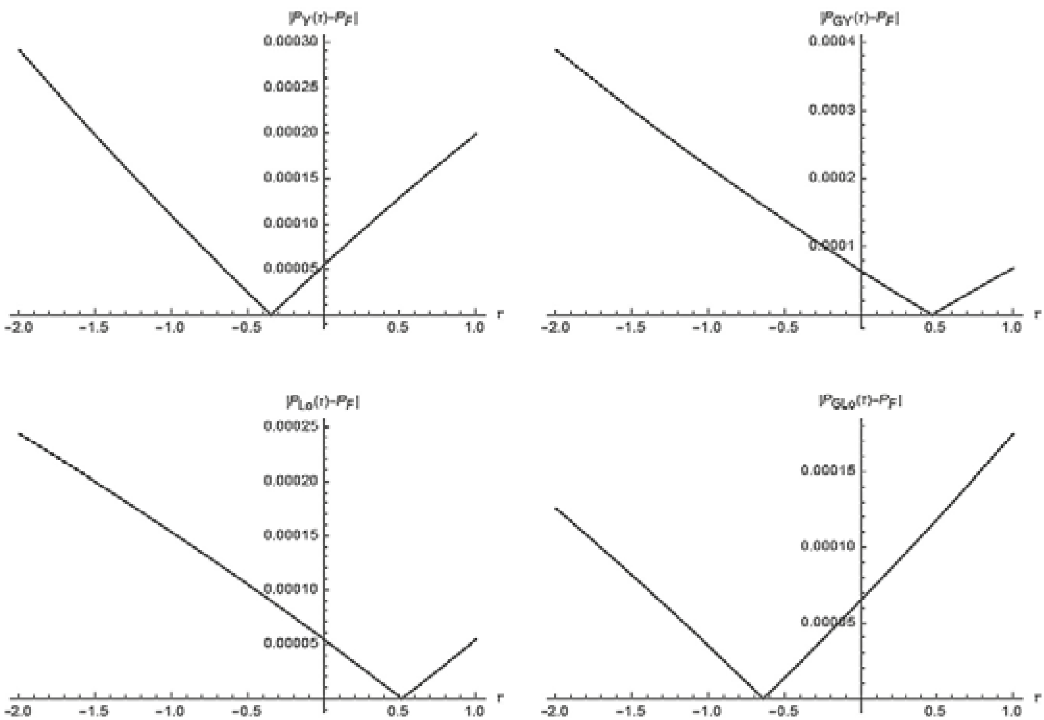
Source: Own construction in Mathematica 11

Table 12 Distance between considered price indices and the Fisher index for the case 1d

τ parameter	P_L-P_F	P_Y-P_F	P_{GY}-P_F	P_{Lo}-P_F	P_{GLo}-P_F
–2.00	0.0000545	–0.0002913	–0.0003892	0.0002437	0.0001259
–1.00	0.0000545	–0.0001089	–0.0002178	0.0001534	0.0000348
–0.75	0.0000545	–0.0000663	–0.0001779	0.0001296	0.0000107
–0.50	0.0000545	–0.0000249	–0.0001391	0.0001051	–0.0000139
–0.25	0.0000545	0.0000154	–0.0001015	0.0000801	–0.0000392
0.25	0.0000545	0.0000924	–0.0000298	0.0000282	–0.0000915
0.50	0.0000545	0.0001291	0.0000044	0.0000014	–0.0001187
0.75	0.0000545	0.0001647	0.0000373	-0.0000262	–0.0001465

Source: Own construction in Mathematica 11

Figure 4 Absolute differences between the Fisher index and the considered price indices as functions of τ (case 1d)



Source: Own construction in Mathematica 11

Case 2a

Table 13 The values of the considered price indices and their distances to the Fisher price index for the case 2a

τ parameter	Laspeyres	Paasche	Fisher	Young	Geo.Young	Lowe	Geo.Lowe
–2.00	1.1389	1.1341	1.1365	1.1380	1.1353	1.1492	1.1463
–1.00	1.1389	1.1341	1.1365	1.1384	1.1357	1.1439	1.1411
–0.75	1.1389	1.1341	1.1365	1.1385	1.1358	1.1426	1.1398
–0.50	1.1389	1.1341	1.1365	1.1386	1.1359	1.1414	1.1386
–0.25	1.1389	1.1341	1.1365	1.1387	1.1360	1.1401	1.1373
0.25	1.1389	1.1341	1.1365	1.1390	1.1363	1.1376	1.1349
0.50	1.1389	1.1341	1.1365	1.1392	1.1364	1.1364	1.1338
0.75	1.1389	1.1341	1.1365	1.1394	1.1366	1.1353	1.1326

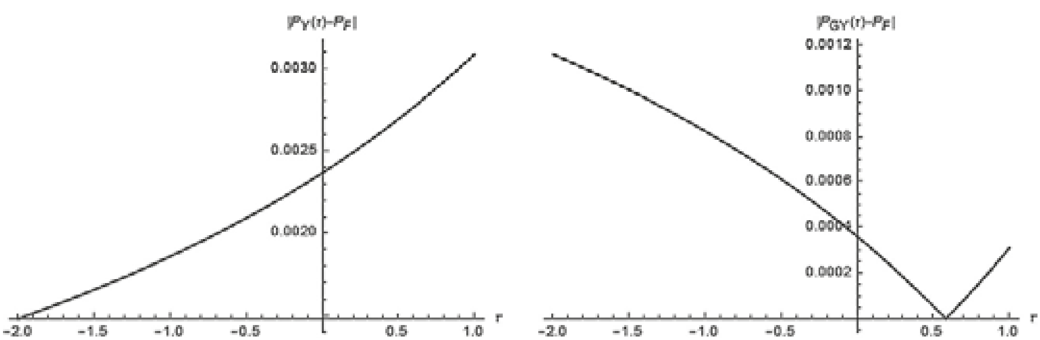
Source: Own construction in Mathematica 11

Table 14 Distance between considered price indices and the Fisher index for the case 2a

τ parameter	P_L-P_F	P_Y-P_F	P_{GY}-P_F	P_{Lo}-P_F	P_{GLo}-P_F
–2.00	0.0023675	0.0014823	–0.0011591	0.0127277	0.0098000
–1.00	0.0023675	0.0018598	–0.0008239	0.0074280	0.0045957
–0.75	0.0023675	0.0019715	–0.0007228	0.0061365	0.0033298
–0.50	0.0023675	0.0020925	–0.0006125	0.0048617	0.0020810
–0.25	0.0023675	0.0022241	–0.0004919	0.0036049	0.0008507
0.25	0.0023675	0.0025239	–0.0002147	0.0011509	–0.0015489
0.50	0.0023675	0.0026947	–0.0000557	–0.0000437	–0.0027158
0.75	0.0023675	0.0028810	0.0001184	–0.0012152	–0.0038594

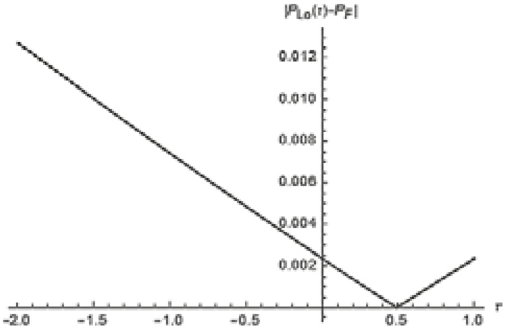
Source: Own construction in Mathematica 11

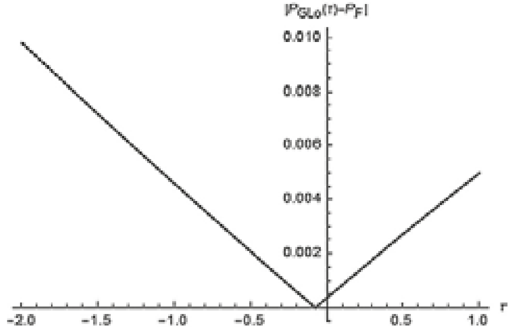
Figure 5 Absolute differences between the Fisher index and the considered price indices as functions of τ (case 2a)



Source: Own construction in Mathematica 11

Figure 5 (continuation)





Source: Own construction in Mathematica 11

Case 2b

Table 15 The values of the considered price indices and their distances to the Fisher price index for the case 2b

τ parameter	Laspeyres	Paasche	Fisher	Young	Geo.Young	Lowe	Geo.Lowe
−2.00	0.9358	0.9353	0.9355	0.935425	0.9350	0.9368	0.9365
−1.00	0.9358	0.9353	0.9355	0.9357	0.9353	0.9363	0.9360
−0.75	0.9358	0.9353	0.9355	0.9357	0.9354	0.9362	0.9359
−0.50	0.9358	0.9353	0.9355	0.9357	0.9354	0.9360	0.9358
−0.25	0.9358	0.9353	0.9355	0.9358	0.9355	0.9359	0.9356
0.25	0.9358	0.9353	0.9355	0.9358	0.9355	0.9357	0.9354
0.50	0.9358	0.9353	0.9355	0.9358	0.9355	0.9355	0.9352
0.75	0.9358	0.9353	0.9355	0.9358	0.9355	0.9354	0.9351

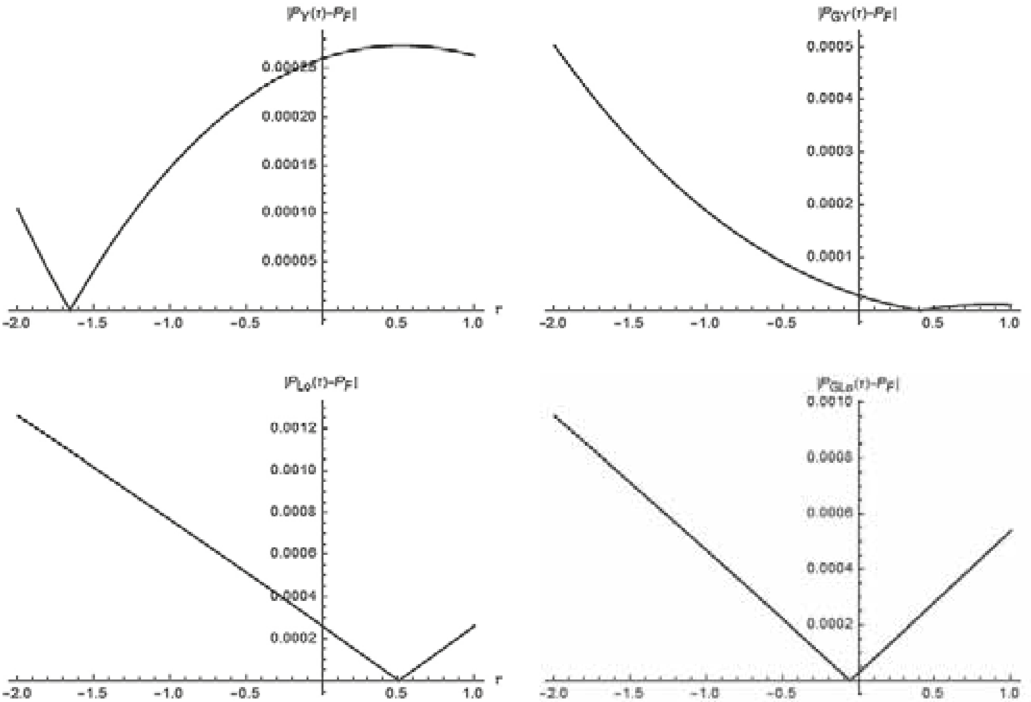
Source: Own construction in Mathematica 11

Table 16 Distance between considered price indices and the Fisher index for the case 2b

τ parameter	P_L-P_F	P_Y-P_F	P_{GY}-P_F	P_{Lo}-P_F	P_{GLo}-P_F
−2.00	0.0002600	−0.0001042	−0.0005026	0.0012598	0.0009538
−1.00	0.0002600	0.0001474	−0.0001885	0.0007666	0.0004693
−0.75	0.0002600	0.0001873	−0.0001355	0.0006412	0.0003461
−0.50	0.0002600	0.0002190	−0.0000915	0.0005150	0.0002220
−0.25	0.0002600	0.0002431	−0.0000561	0.0003879	0.0000971
0.25	0.0002600	0.0002700	−0.0000087	0.0001312	−0.0001552
0.50	0.0002600	0.0002736	0.0000042	0.0000017	−0.0002826
0.75	0.0002600	0.0002712	0.0000104	−0.0001287	−0.0004108

Source: Own construction in Mathematica 11

Figure 6 Absolute differences between the Fisher index and the considered price indices as functions of *τ* (case 2b)



Source: Own construction in Mathematica 11

Case 2c

Table 17 The values of the considered price indices and their distances to the Fisher price index for the case 2c

τ parameter	Laspeyres	Paasche	Fisher	Young	Geo.Young	Lowe	Geo.Lowe
–2.00	1.00942	1.0053	1.0074	1.0036	0.9998	1.0185	1.0148
–1.00	1.00942	1.0053	1.0074	1.0062	1.0024	1.0138	1.0100
–0.75	1.00942	1.0053	1.0074	1.0069	1.0031	1.0127	1.0088
–0.50	1.00942	1.0053	1.0074	1.0077	1.0039	1.0116	1.0077
–0.25	1.00942	1.0053	1.0074	1.0085	1.0046	1.0105	1.0066
0.25	1.00942	1.0053	1.0074	1.0104	1.0064	1.0084	1.0044
0.50	1.00942	1.0053	1.0074	1.0114	1.0073	1.0073	1.0034
0.75	1.00942	1.0053	1.0074	1.0124	1.0083	1.0063	1.0023

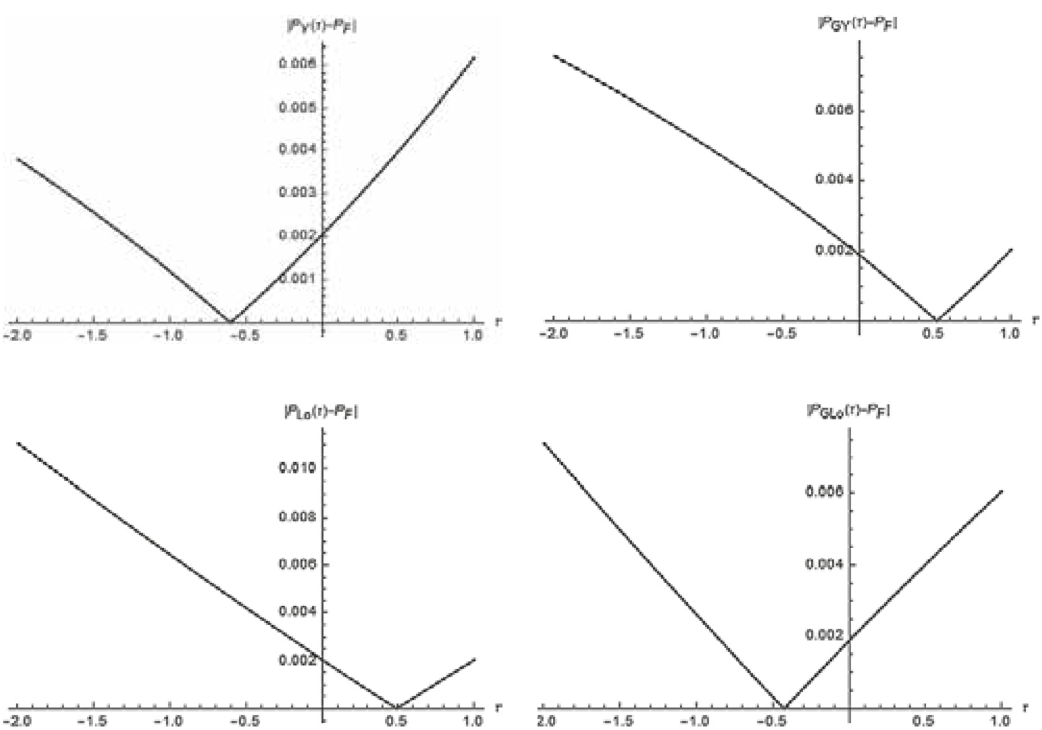
Source: Own construction in Mathematica 11

Table 18 Distance between considered price indices and the Fisher index for the case 2c

τ parameter	P_L-P_F	P_Y-P_F	P_{GY}-P_F	P_{Lo}-P_F	P_{GLo}-P_F
–2.00	0.0020476	–0.0038098	–0.0075550	0.0110956	0.0073880
–1.00	0.0020476	–0.0011955	–0.0049864	0.0064462	0.0026152
–0.75	0.0020476	–0.0004550	–0.0042735	0.0053202	0.0014618
–0.50	0.0020476	0.0003298	–0.0035221	0.0042111	0.0003266
–0.25	0.0020476	0.0011627	–0.0027274	0.0031200	–0.0007894
0.25	0.0020476	0.0029876	–0.0009925	0.0009945	–0.0029610
0.50	0.0020476	0.0039858	–0.0000451	–0.0000385	–0.0040155
0.75	0.0020476	0.0050447	0.0009596	–0.0010513	–0.0050487

Source: Own construction in Mathematica 11

Figure 7 Absolute differences between the Fisher index and the considered price indices as functions of τ (case 2c)



Source: Own construction in Mathematica 11

Case 2d

Table 19 The values of the considered price indices and their distances to the Fisher price index for the case 2d

τ parameter	Laspeyres	Paasche	Fisher	Young	Geo.Young	Lowe	Geo.Lowe
–2.00	1.02515	1.0250	1.0251	1.0249	1.0248	1.0254	1.0253
–1.00	1.02515	1.0250	1.0251	1.0250	1.0249	1.0253	1.0251
–0.75	1.02515	1.0250	1.0251	1.0251	1.0250	1.0252	1.0251
–0.50	1.02515	1.0250	1.0251	1.0251	1.0250	1.0252	1.0251
–0.25	1.02515	1.0250	1.0251	1.0251	1.0250	1.0252	1.0251
0.25	1.02515	1.0250	1.0251	1.0252	1.0251	1.0251	1.0250
0.50	1.02515	1.0250	1.0251	1.0252	1.0251	1.0251	1.0250
0.75	1.02515	1.0250	1.0251	1.0253	1.0251	1.0251	1.0250

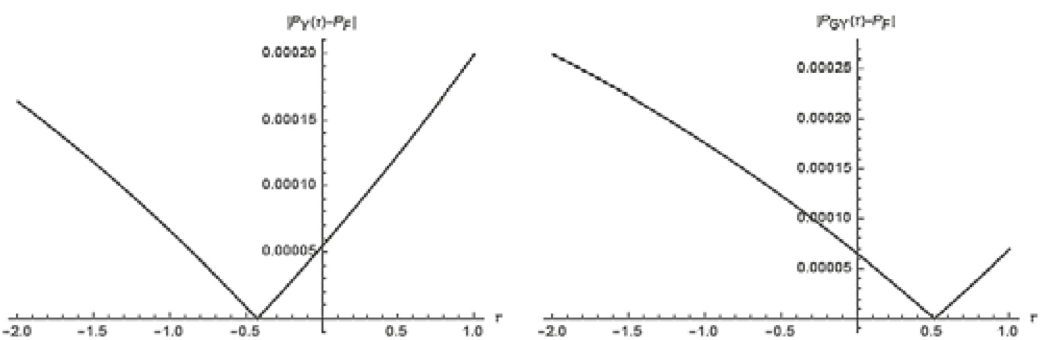
Source: Own construction in Mathematica 11

Table 20 Distance between considered price indices and the Fisher index for the case 2d

τ parameter	P_L-P_F	P_Y-P_F	P_{GY}-P_F	P_{Lo}-P_F	P_{GLo}-P_F
–2.00	0.0000545	–0.0001638	–0.0002645	0.0002751	0.0001586
–1.00	0.0000545	–0.0000661	–0.0001759	0.0001646	0.0000464
–0.75	0.0000545	–0.0000382	–0.0001503	0.0001370	0.0000184
–0.50	0.0000545	–0.0000087	–0.0001233	0.0001095	–0.0000095
–0.25	0.0000545	0.0000221	0.0000949	0.0000820	–0.0000373
0.25	0.0000545	0.0000884	–0.0000337	0.0000271	–0.0000927
0.50	0.0000545	0.0001237	–0.0000009	–0.0000002	–0.0001203
0.75	0.0000545	0.0001606	0.0000333	–0.0000274	–0.0001477

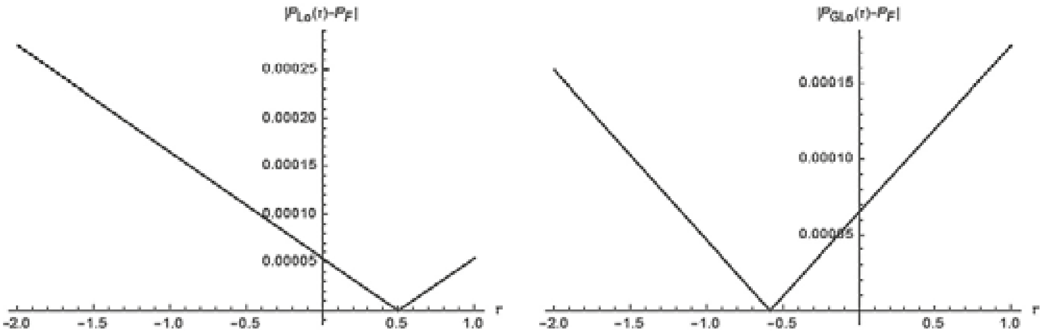
Source: Own construction in Mathematica 11

Figure 8 Absolute differences between the Fisher index and the considered price indices as functions of τ (case 2d)



Source: Own construction in Mathematica 11

Figure 8 (continuation)



Source: Own construction in Mathematica 11

6 EMPIRICAL STUDY

In this section, we wish to verify the level of bias altering the above-mentioned indices. We collect data from the COICOP 3 and 4 level. We consider the following groups of goods and services from the HICP basket:

- Food,
- Alcoholic beverages,
- Audio-visual, photographic and information processing equipment,
- Newspapers, books and stationery.

We compare results of both mean and summed up substitution bias calculated for years 2006–2018 for Poland, Czech Republic, Hungary, Slovakia, United Kingdom, France, Germany, and the UE benchmark. We take $\tau = -1$ for calculations of the Young and the Lowe indices. These results are presented in Tables 21–28.

In all considered groups of goods mean value for substitution bias was the smallest in the case of the Laspeyres index. Even though geometric versions of both the Lowe and the Young indices gave much better results than their arithmetic counterparts, the substitution bias was still considerately bigger than for the Laspeyres index. This is partly the effect of year-to-year update of consumer baskets in both CPI and HICP indices as well as that the HICP index by definition already tries to reduce substitution bias.

Table 21 Mean values of differences between the considered indices and Fisher index for "food" category

Country	P_L-P_F	P_{Lo}-P_F	P_Y-P_F	P_{GLo}-P_F	P_{GY}-P_F
European Union	0.00008	0.00065	0.00070	0.00057	0.00062
Czechia	0.00060	0.00314	0.00257	0.00316	0.00219
Germany	0.00019	0.00093	0.00099	0.00076	0.00082
France	0.00009	0.00057	0.00066	0.00054	0.00060
Hungary	0.00103	0.00160	0.00229	0.00104	0.00160
Poland	0.00041	0.00152	0.00166	0.00116	0.00126
Slovakia	0.00036	0.00116	0.00169	0.00067	0.00123
United Kingdom	0.00013	0.00052	0.00064	0.00046	0.00059

Source: Own construction in Mathematica 11

Table 22 Summed up values of differences between the considered indices and Fisher index for "food" category

Country	P_L-P_F	P_{Lo}-P_F	P_Y-P_F	P_{GLo}-P_F	P_{GY}-P_F
European Union	0.00080	0.00831	0.00896	0.00688	0.00753
Czechia	0.00721	–0.01914	–0.01310	–0.02938	–0.02334
Germany	0.00177	0.01161	0.01187	0.00834	0.00864
France	0.00105	0.00681	0.00799	0.00549	0.00668
Hungary	0.01339	0.01626	0.02974	0.00248	0.01438
Poland	0.00203	0.01757	0.01914	0.01188	0.01339
Slovakia	0.00203	0.01460	0.02077	0.00561	0.01111
United Kingdom	0.00131	0.00658	0.00791	0.00498	0.00630

Source: Own construction in Mathematica 11

Table 23 Mean values of differences between the considered indices and Fisher index for "alcoholic beverages" category

Country	P_L-P_F	P_{Lo}-P_F	P_Y-P_F	P_{GLo}-P_F	P_{GY}-P_F
European Union	0.00001	0.00004	0.00004	0.00004	0.00005
Czechia	0.00025	0.00060	0.00060	0.00060	0.00060
Germany	0.00006	0.00033	0.00032	0.00031	0.00032
France	0.00015	0.00054	0.00053	0.00057	0.00056
Hungary	0.00052	0.00200	0.00185	0.00192	0.00176
Poland	0.00006	0.00021	0.00018	0.00023	0.00020
Slovakia	0.00016	0.00040	0.00041	0.00043	0.00045
United Kingdom	0.00012	0.00056	0.00053	0.00054	0.00051

Source: Own construction in Mathematica 11

Table 24 Summed up values of differences between the considered indices and Fisher index for "alcoholic beverages" category

Country	P_L-P_F	P_{Lo}-P_F	P_Y-P_F	P_{GLo}-P_F	P_{GY}-P_F
European Union	0.00011	0.00020	0.00017	0.00011	0.00008
Czechia	0.00121	0.00140	0.00208	0.00021	0.00088
Germany	–0.00036	0.00305	0.00273	0.00273	0.00242
France	0.00120	0.00322	0.00342	0.00223	0.00242
Hungary	0.00501	–0.00935	–0.00784	–0.01354	–0.01202
Poland	0.00059	–0.00109	–0.00134	–0.00161	–0.00186
Slovakia	0.00103	–0.00159	–0.00179	–0.00266	–0.00286
United Kingdom	0.00145	–0.00168	–0.00169	–0.00259	–0.00260

Source: Own construction in Mathematica 11

Table 25 Mean values of differences between the considered indices and Fisher index for "newspapers, books, and stationery" category

Country	P_L-P_F	P_{Lo}-P_F	P_Y-P_F	P_{GLo}-P_F	P_{GY}-P_F
European Union	0.00012	0.00054	0.00040	0.00048	0.00037
Czechia	0.00049	0.00126	0.00114	0.00122	0.00110
Germany	0.00018	0.00083	0.00059	0.00075	0.00056
France	0.00011	0.00028	0.00019	0.00022	0.00016
Hungary	0.00212	0.00566	0.00404	0.00452	0.00305
Poland	0.00074	0.00148	0.00181	0.00144	0.00204
Slovakia	0.00037	0.00130	0.00110	0.00122	0.00101
United Kingdom	0.00045	0.00113	0.00095	0.00116	0.00098

Source: Own construction in Mathematica 11

Table 26 Summed up values of differences between the considered indices and Fisher index for "newspapers, books, and stationery" category

Country	P_L-P_F	P_{Lo}-P_F	P_Y-P_F	P_{GLo}-P_F	P_{GY}-P_F
European Union	0.00144	0.00062	–0.00042	–0.00028	–0.00132
Czechia	0.00361	0.00187	0.00189	0.00065	0.00067
Germany	0.00161	–0.00147	–0.00206	–0.00323	–0.00383
France	0.00081	0.00341	0.00198	0.00243	0.00101
Hungary	0.02730	0.07291	0.04700	0.05707	0.03065
Poland	0.00355	0.00157	–0.00735	–0.00651	–0.01541
Slovakia	0.00419	0.00210	0.00317	–0.00054	0.00054
United Kingdom	0.00050	–0.00959	–0.00898	–0.01208	–0.01149

Source: Own construction in Mathematica 11

Table 27 Mean values of differences between the considered indices and Fisher index for "audio-visual, photographic and information processing equipment" category

Country	P_L-P_F	P_{Lo}-P_F	P_Y-P_F	P_{GLo}-P_F	P_{GY}-P_F
European Union	0.00115	0.00449	0.00223	0.00389	0.00146
Czechia	0.00132	0.00610	0.00421	0.00594	0.00418
Germany	0.00112	0.00275	0.00261	0.00241	0.00247
France	0.00144	0.00488	0.00444	0.00443	0.00420
Hungary	0.00127	0.00392	0.00492	0.00366	0.00464
Poland	0.00124	0.00327	0.00352	0.00260	0.00293
Slovakia	0.00161	0.00540	0.00508	0.00495	0.00462
United Kingdom	0.00285	0.01247	0.00543	0.01091	0.00384

Source: Own construction in Mathematica 11

Table 28 Summed up values of differences between the considered indices and Fisher index for "audio-visual, photographic and information processing equipment" category

Country	P_L-P_F	P_{Lo}-P_F	P_Y-P_F	P_{GLo}-P_F	P_{GY}-P_F
European Union	0.01448	0.01608	0.02812	0.00489	0.01652
Czechia	0.01390	0.00429	0.01990	–0.01028	0.00494
Germany	0.01358	0.02909	0.02989	0.02053	0.02126
France	0.00931	0.00911	0.01397	–0.00406	0.00038
Hungary	0.01205	0.04547	0.05654	0.04081	0.05154
Poland	0.01329	0.03267	0.03319	0.02279	0.02319
Slovakia	0.01440	0.03305	0.03880	0.02078	0.02658
United Kingdom	0.03699	0.02287	0.05110	–0.00700	0.01923

Source: Own construction in Mathematica 11

However, it is worth mentioning that aggregated bias from the period 2006–2018 was bigger for the Laspeyres index in some cases, especially for "audio-visual, photographic, and information processing equipment" and "newspapers, books, and stationery". Due to the fact that the Lowe and the Young indices bias direction is more unpredictable than in the case of the Laspeyres index which regularly overstates inflation, in some cases for a long period of time they might be a better option.

CONCLUSION

In the majority of considered cases the Young index gives better approximation of the Fisher Index than the Laspeyres index, for the **τ in the range** [–1; –0.25] and worse in range [0.25; 1], which still makes him the most reliable, as usually, statisticians use periods prior to base period. However, for linear price decrease (1b) the Young index gave the opposite results – it was biased in range [–1; –0.25] and less biased in range [0.25; 1]. In this case, the Geometric Young index gives much better results, even though it was biased in most of the other considered simulations.

In every simulation the Lowe index gave the worse results for τ in the range [–1; –0.25] and better in [0.25; 1], thus making it unreliable for statistical purposes if we wish to use data from the previous time periods. However, its geometric counterpart gave better results, especially for the fourth basket with an inflation rate close to 2.5%. In both, linear and exponential cases, it gave better results than the Young index, thus becoming an interesting alternative for measuring inflation in stable conditions in developed countries.

For empirical data in different groups of goods, the average bias for both Young and Lowe indices was bigger than for the Laspeyres index. However, in the case of some groups of wares, the aggregated bias of the Young and the Lowe indices was much smaller, thus making it an interesting alternative for inflation measurement in the long-term.

References

ARMKNECHT, P. AND SILVER, M. Post- Laspeyres: the Case for a New Formula for Compiling Consumer Price Index. *IMF Working Paper*, 2012, WP/12/105.

BIAŁEK, J. Approximation of the Fisher price index by using Lowe, Young, and AG Mean indices. *Communications in Statistics – Simulation and Computation*, 2017, 46, 8, pp. 6454–6467.

BREUER, C. AND VON DER LIPPE, P. *Problems of operationalizing the concept of a "cost of living index"*. MPRA Paper No. 32902, 2011.
CAMBA-MENDEZ, G. et al. *Measurement Issues in European Consumer Price Indices and the Conceptual Framework of the HICP*. European Central Bank, 2002.
DIEWERT, W. E. The Theory of the Cost-of-Living Index and the Measurement of Welfare Change. *Contributions to Economic Analysis*, 1990, Vol. 196, pp. 79–147.
DIEWERT, W. E. AND VON DER LIPPE, P. Introduction to the Special Issue on Index Number Theory and Price Statistics. *Jahrbücher für Nationalökonomie und Statistik*, 2010, Bd. 6/230, pp. 660–672 (661).
DORFMAN, A. AND LENT, J. Using a Weighted Average of Base Period Price to Approximate a Superlative Index. *Journal of Official Statistics,* 2009, Vol. 25, No. 1 pp. 139–149.
EUROSTAT. *HICP data* [online]. <https://ec.europa.eu>.
FISHER, R. A. On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Society*, 1922, 222, pp. 309–368.
HAŁKA, A. AND LESZCZYŃSKA, A. Wady i zalety wskaźnika cen towarów i usług konsumpcyjnych – szacunki obciążenia dla Polski. *Gospodarka Narodowa*, 2011 (September), pp. 60–61.
HUANG, N. et al. *Choice of index number formula and the upper-level substitution bias in the Canadian CPI*. Statistics Canada, Consumer Prices Division, 14th Ottawa Group Meeting Tokyo, Japan, 20–22 May, 2015.
INTERNATIONAL LABOR ORGANIZATION. *An Introduction to consumer price index methodology* [online]. ILO. [cit. 20.5.2019] <http://www.ilo.org/public/english/bureau/stat/download/cpi/ch17.pdf>.
INTERNATIONAL LABOR ORGANIZATION. *The Economic Approach to index number Theory: the single household case* [online]. ILO. [cit. 16.5.2019] <http://www.ilo.org/public/english/bureau/stat/download/cpi/ch17.pdf>.
INTERNATIONAL MONETARY FUND. *Axiomatic and Stochastic Approaches to Index Number Theory* [online]. IMF. [cit. 10.6.2019] <https://www.imf.org/external/np/sta/tegeipi/ch17.pdf>.
KARSAULIDZE, L. *Possible Reasons of Bias in Estimating the Cost of Living Index by the CPI.* Switzerland, 2018.
LLOYD, P. J. Substitution Effects and Biases in Nontrue Price Indices. *The American Economic Review*, 1975, pp. 301–313.
MOULTON, B. *Constant Elasticity Cost of Living Index in Share-Relative Form*. Unpublished U.S. Bureau of Labor Statistics manuscript, 1996.
VEBLEN, T. *The Theory of the Leisure Class*, Colombia University [online]. New York: Macmillan, 1899. [cit. 15.6.2019] <http://moglen.law.columbia.edu/LCS/theoryleisureclass.pdf>.
VON DER LIPPE, P. *Index Theory and Price Statistics*. Frankfurt, Germany: Peter Lang, 2007.

Remarks on Price Index Methods for the CPI Measurement Using Scanner Data

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

Abstract

Scanner data are a quite new data source for statistical agencies and the availability of electronic sales data for the calculation of the Consumer Price Index (CPI) has increased over the past 16 years. Scanner data can be obtained from a wide variety of retailers (supermarkets, home electronics, Internet shops, etc.) and provide information at the level of the barcode, i.e. the Global Trade Item Number (GTIN, formerly known as the EAN code). One of new challenges connected with scanner data is the choice of the index formula which should be able to reduce the chain drift bias and the substitution bias. In this paper, we compare several price index methods for CPI calculations based on scanner data. In particular, we consider bilateral index methods with chained versions of direct weighted and unweighted indices, and also selected multilateral index methods, i.e. the quality adjusted unit value method (QU method) and its special case (the Geary-Khamis method), the augmented Lehr method, the so called "real time index", the GEKS method and the CCDI method. We consider different weighting schemes in quantity weights on the price index. We compare all these methods using a real scanner data set obtained from one supermarket chain.. The main aim of the paper is to show how big differences among bilateral and multilateral indices may rise while using real scanner data sets. In particular our results lead to the conclusion that the choice of the multilateral formula and the weighting scheme does matter in inflation measurement. It is shown that differences between values of all discussed formulas may exceed several percentage points even in the case of only one homogeneous group of products.[2]

Keywords

Scanner data, Consumer Price Index, superlative indices, elementary indices, chain indices, QU-GK index, Geary-Khamis method, real time index, GEKS, bilateral indices, multilateral indices

JEL code

C43, E31

INTRODUCTION

Scanner data mean transaction data that specify turnover and numbers of items sold by GTIN (barcode, formerly known as the EAN code). Scanner data have numerous advantages compared to traditional survey data collection because such data sets are much bigger than traditional ones and they contain complete transaction information, i.e. information about prices and quantities.

[1] Department of Statistical Methods, University of Lodz, 90-255 Lodz, 3/5 POW Street, Poland. E-mail: jacek.bialek@uni.lodz.pl. Central Statistical Office in Poland, Department of Trade and Services, Warsaw, Poland.

[2] This is a modified and improved version of the paper titled "Comparison of Price Index Methods for the CPI Measurement Using Scanner Data" which was presented during the 16th *Meeting of the Ottawa Group on Price Indices,* Rio de Janeiro, 8–10 May 2019.

In other words, scanner data contain expenditure information at the item level (i.e. at the barcode or the GTIN level), which makes it possible to use expenditure shares of items as weights for calculating price indices at the lowest (elementary) level of data aggregation.

Scanner data from two supermarkets were introduced in the Dutch CPI in 2002 and, in January 2010, the number of supermarkets providing the scanner data was extended to six. The Dutch CPI was re-designed (de Haan, 2006; Van der Grient and de Haan, 2010; de Haan and Van der Grient, 2011). In 2017, scanner data of ten supermarkets chains were used and at present surveys are not carried out anymore for supermarkets, i.e. scanner data from other retailers (for instance, from do-it-yourself stores or from travel agencies) are used in the Dutch CPI (Chessa, 2015). Until 2015, four EU countries were using scanner data (the Netherlands, Norway, Sweden, and Switzerland). The number of countries that make use of scanner data in their CPI has been growing, i.e. in April 2016, the number of EU countries increased to seven (Belgium, Denmark and Iceland started to use such data sets) and at present, some of national statistical institutes (NSIs) consider starting to use scanner data. Some other countries consider using scanner data in their CPI calculation in the nearest future (or have just started using it), for instance: the French National Statistical Institute (INSEE) launched in 2010 a pilot project in order to get some insights into the suitability of these data for CPI purposes, the Statistics Portugal was awarded in 2011 a Eurostat grant to undertake the initial research on the exploitation of scanner data, in Luxembourg, collaboration was put in place with several retailers who agreed to transmit every month their data to the IT system (STATEC) and scanner data have been introduced in the regular production from January 2018. In January 2018, in Poland, the project titled "INSTATCENY" began and its main aim is to create the new methodology of CPI measurement based on data from different (traditional and untraditional) sources, including scanner data and web-scraped data. In 2017, the Eurostat provided *Practical Guide for Processing Supermarket Scanner data,* which is commonly available on website: <*https://ec.europa.eu/eurostat/web/hicp/overview*>. In the above-mentioned guide, we can read: "This guide describes the situation in 2017. It will need to be updated as the use of scanner data develops and broadens". In fact, the methodology for CPI (or HICP) construction using scanner data has strongly evolved over the last few years (see for instance: Ivancic et. al., 2011; Krsnich, 2014; Griffioen and Bosch, 2016; de Haan et al., 2016; Chessa and Griffioen, 2016; Chessa, 2017; Diewert and Fox, 2017). One of new challenges connected with scanner data is the choice of the index formula which should be able to reduce the chain drift bias and the substitution bias.

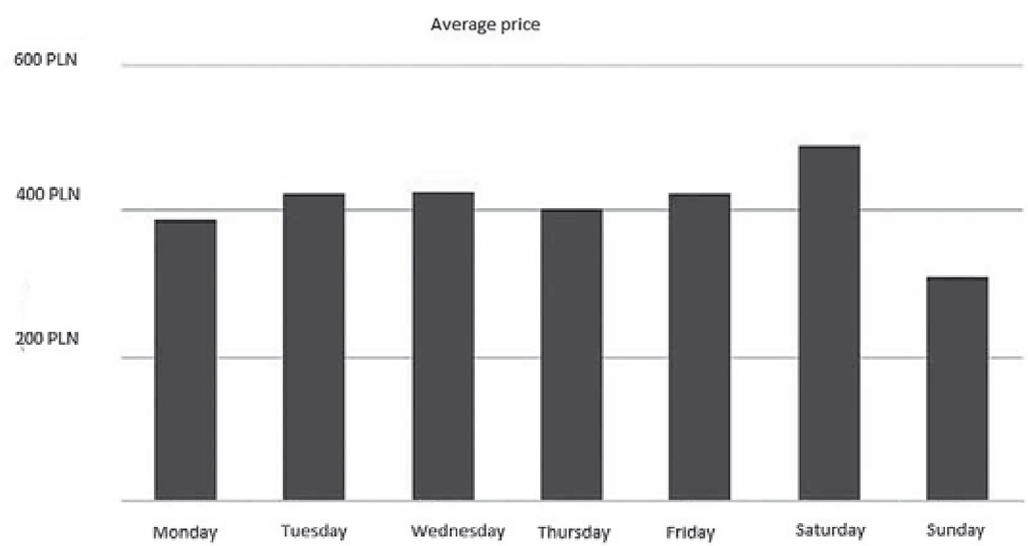
In this paper, we compare several price index methods for CPI calculations based on scanner data. The main aim of the paper is to show how big differences among bilateral and multilateral indices may rise while using real scanner data sets. In particular it is shown that the choice of the multilateral formula and the weighting scheme does matter in inflation measurement. The paper is organised as follows: Section 1 presents main advantages, disadvantages and challenges connected with using scanner data. Section 2 describes a selected bilateral and multilateral index method which can be used in the case of scanner data and this Section also discusses updating and weighting problem connected with multilateral methods; Section 3 proposes some price index modifications; Section 4 presents the results from our simulation study and examines the influence of the price and quantity behaviour on differences between the discussed index methods; Section 5 continues the comparison of bilateral and multilateral methods; it presents the empirical study based on real scanner data sets obtained from one supermarket and the e-commerce platform *allegro.pl*; last section lists the main conclusions.

1 SCANNER DATA: ADVANTAGES, DISADVANTAGES AND CHALLENGES

One of the main advantages of using scanner data is the fact that these data sets allow for the level of elementary aggregate to be taken down to lower levels, as the information about prices and quantities (thus also about weights) is available. Scanner data sets are huge and may provide some additional information about products (such as the following attributes: size, colour, package quantity, etc.). These attributes may be useful in aggregating items into homogeneous groups. Besides, obtaining scanner data is much cheaper than obtaining CPI data in the traditional manner. In the Eurostat's *Practical Guide for Processing Supermarket Scanner data* from 2017, we can read (page 9): "In the traditional price collection, price collectors have to trust intuition and common sense and it may happen that prices are collected as long as the item is available even though it is no longer representative. In scanner data the representativeness is guaranteed".

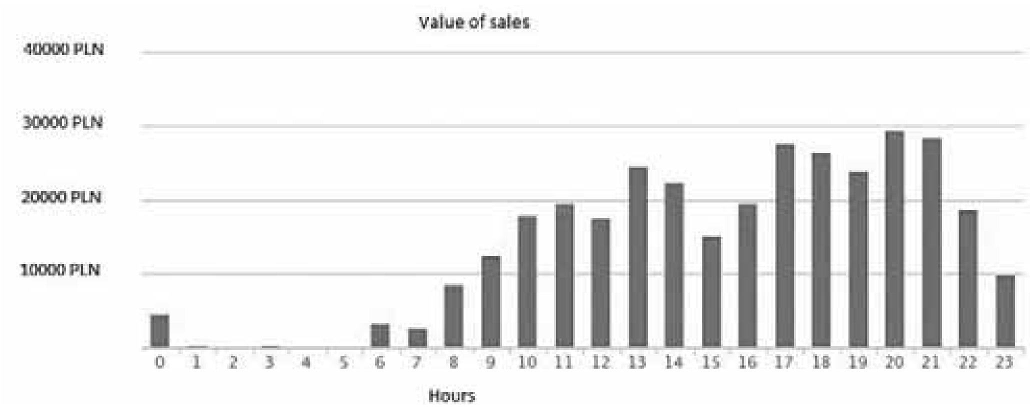
To list disadvantages of using scanner data, we should start with the substantial dependence of the NSI on a retailer. In fact, in the traditional CPI measurement, a price collector must only receive from a given retailer the permission to visit the outlet. In the case of scanner data, there must be a legal contract between the NSI and a given retailer which allows the NSI to fully control and monitor the data. The second disadvantage (or rather a new challenge) is the fact that methodology for sampling of scanner data is still poorly developed, i.e. there are open questions about sampling of regions (if retailers use regional prices) or sampling of outlets (if the outlets differ from one another with respect to opening hours or offered goods). Also the choice of the time interval for aggregating of scanner data may be problematic, since many product prices change periodically but the price cycle may equal to a quarter, a month, a week or even a day. In fact, observing scanner data, it is easy to confirm the known fact that prices are often higher on Saturday (see Figure 1), and each day the value of sales (see Figure 2), the number of transactions (see Figure 3) and the price (see Figure 4) are the highest in the afternoon or in the evening. Moreover, the distribution of the expenditure on a given product may strongly depend on the day, for instance, the expenditure fluctuations on Monday and on Friday may differ from each other significantly (see Figure 5). The same remark could be repeated for prices of products (see Figure 6). All figures which confirm our above-mentioned remarks, i.e. Figures 1–6, are obtained for the homogeneous group of child safety seats (sample EANs: 5902581655226, 5902533903429, or 5902581652850) and the analysed scanner data set comes from the net portal *allegro.pl* (the *TradeWatch* tool) and concerns the time interval: 22.11.2015–16.12.2018.

Figure 1 Daily average price of child safety seats



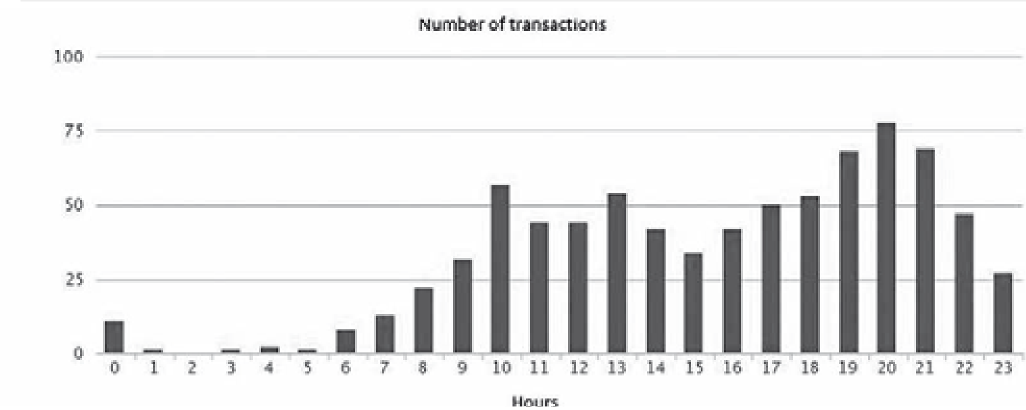
Source: <*TradeWatch.pl*>

Figure 2 Average value of sales of child safety seats depending on hours



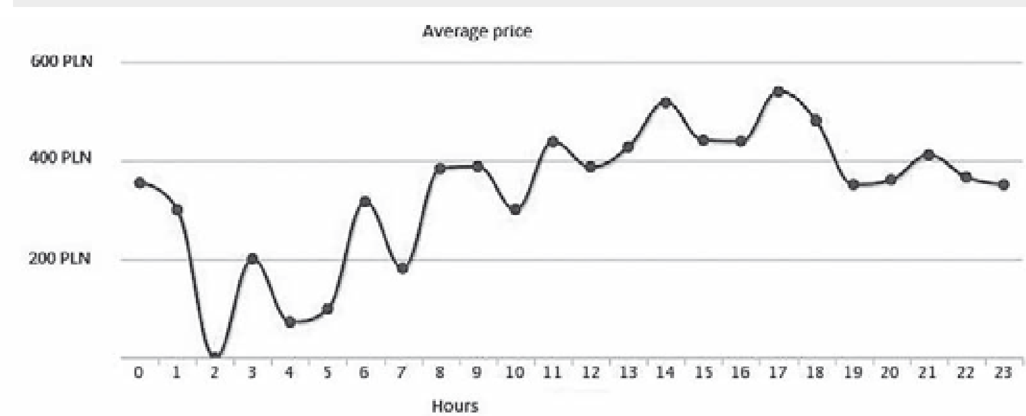
Source: *<TradeWatch.pl>*

Figure 3 Average number of transactions concerning child safety seats depending on hours



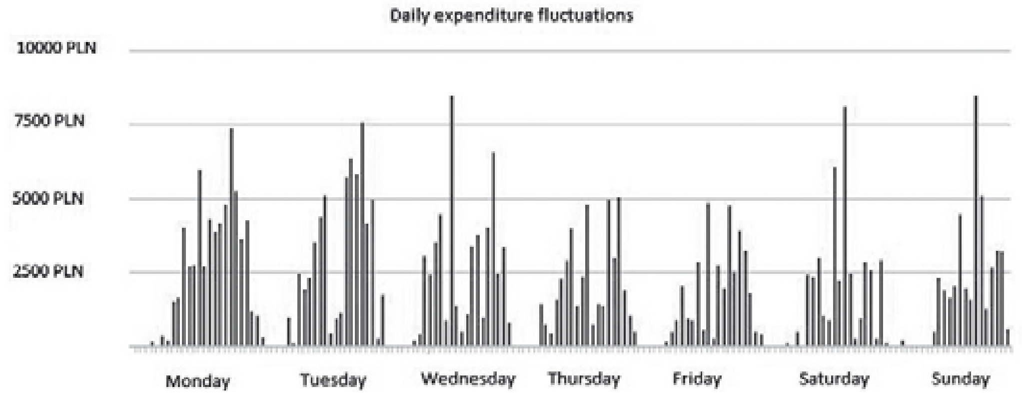
Source: *<TradeWatch.pl>*

Figure 4 Average price of child safety seats depending on hours



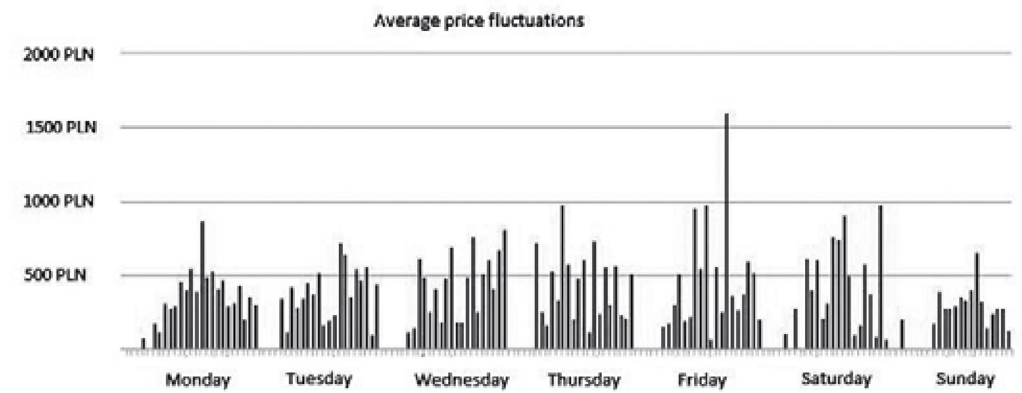
Source: *<TradeWatch.pl>*

Figure 5 Fluctuations of the average expenditure on child safety seats depending on days



Source: <*TradeWatch.pl*>

Figure 6 Fluctuations of the average price of child safety seats depending on days



Source: <*TradeWatch.pl*>

The first challenge connected with scanner data concerns item codes. As it was mentioned above, the GTIN (formerly known as the EAN) is the name currently used for coding in the case of scanner data. Nevertheless, the following codes may also be used: price look-up (PLU) and stock-keeping units (SKUs). PLU codes are shorter than GTINs and SKUs can be slightly more generic than GTINs. In practice, when NSIs use scanned data from different retailers with different code systems, some problems with identifying products may rise. Moreover, scanner data may contain data on transactions between the retailer and other business, which should be verified and detected (such transactions should be excluded from CPI calculations). The next challenge is detecting items which were returned within the given period after the purchase. Since typically, 10 000–25 000 item codes are used in the supermarket, a huge challenge is to create the appropriate, preferably automatic (or at least almost fully automatic) IT system which is able to go through with the above-mentioned detections and which takes into consideration seasonal goods, replacements, as well as disappearing and appearing item codes in the sample. Finally, one of new challenges connected with scanner data is the choice of the index formula which should be able to reduce the chain drift bias and the substitution bias. In the paper, we focus on the choice of the price index formula.

2 INDEX METHODS FOR CPI CALCULATIONS USING SCANNER DATA

Most statistical agencies use bilateral index numbers in the CPI measurement, i.e. they use indices which compare prices and quantities of a group of commodities from the current period with the corresponding prices and quantities from a base (fixed) period. In multilateral methods, we collect information about prices and quantities of a group of commodities from T periods and next we calculate a sequence of price indices for these T periods. Although Ivancic, Diewert and Fox (2011) have suggested that the use of multilateral indices in the scanner data case can solve the chain drift problem, most statistical agencies using scanner data still make use of the monthly chained Jevons index (Chessa et al., 2017). Since the elementary Jevons price index belongs to bilateral (direct) index methods, we start our description of possible methods with these methods. Following Chessa et al. (2017), let us denote the sets of homogeneous products belonging to the same product group in months 0 and t by G_0 and G_t respectively, and let $G_{0,t}$ denote the set of matched products in both moments 0 and t. A product may refer to a single item (GTIN) or to a sub-group of items (GTINs) having the same characteristics, and thus being in the same homogeneity group. In the next part of the paper, we consider the second scenario, i.e. a homogeneous group of different GTINs but having identical characteristics. We also consider a month as a time period over which scanner data are aggregated. In fact, one month is the longest interval among time intervals recommended by Eurostat for the scanner data aggregation (see *Practical Guide for Processing Supermarket Scanner data,* 2017, page 13) although, the same document on the same page states: "Most commonly, scanner data are collected weekly, i.e. all transactions taking place during a week are aggregated".

2.1 Bilateral index methods

2.1.1 Unweighted formulas

A recommendation of the European Commission concerning the choice of the elementary formula at the lowest level of data aggregation can be found on website: <*http://www.ilo.org/public/english/bureau/stat/download/cpi/corrections/annex1.pdf*> and it is as follows: "For the HICPs the ratio of geometric mean prices or the ratio of arithmetic mean prices are the two formulae which should be used within elementary aggregates. The arithmetic mean of price relatives may only be applied in exceptional cases and where it can be shown that it is comparable". In other words, if expenditure information is not available, the European Commission recommends the Jevons (1865) price index (see also Diewert, 2012; or Levell, 2015), which can be written as follows:

$$P_J^{0,t} = \prod_{i \in G_{0,t}} \left(\frac{p_i^t}{p_i^0}\right)^{\frac{1}{N_{0,t}}}, \tag{1}$$

where p_i^τ denotes the price of the i-th product at the time $\tau \in \{0, t\}$ and $N_{0,t} = card\ G_{0,t}$. On the other hand, the same recommendation takes also into consideration ("in exceptional cases") the Carli (1804) price index, which can be written as follows:

$$P_C^{0,t} = \frac{1}{N_{0,t}} \sum_{i \in G_{0,t}} \frac{p_i^t}{p_i^0}, \tag{2}$$

In our research, we consider only the first Formula (1) together with its monthly chained version which is denoted here by $P_{CH-J}^{0,t}$.

2.1.2 Weighted formulas

Since scanner data contain information about the expenditure, it is possible in their case to calculate weighted bilateral indices. *Superlative* price indices, firstly proposed by Diewert (1976), are the most recommended index formulas for the scanner data case (as base formulas). Following Chessa et al. (2017), we consider the Törnqvist (1936) price index, which is given by:

$$P_T^{0,t} = \prod_{i \in G_{0,t}} \left(\frac{p_i^t}{p_i^0}\right)^{\frac{s_i^0 + s_i^t}{2}}, \tag{3}$$

where s_i^0 and s_i^t denote the expenditure shares of matched products in months 0 and t.

Other commonly known superlative price indices are the Fisher price index (1922) and the Walsh price index (1901). Their formulas, denoted by $P_F^{0,t}$ and by $P_W^{0,t}$ respectively, can be written as follows:

$$P_F^{0,t} = \sqrt{P_{La}^{0,t} \cdot P_{Pa}^{0,t}}, \tag{4}$$

$$I_W^{0,t} = \frac{\sum_{i \in G_{0,t}} p_i^t \cdot \sqrt{q_i^0 q_i^t}}{\sum_{i \in G_{0,t}} p_i^0 \cdot \sqrt{q_i^0 q_i^t}}, \tag{5}$$

where q_i^0 and q_i^t denote quantities of matched products in months 0 and t, $P_{La}^{0,t}$ and $P_{Pa}^{0,t}$ denote the Laspeyres price index (1864) and the Paasche price index (1874) respectively (see Section 2.2.1). In the next part of the paper only the Fisher and the Törnqvist price indices are taken into consideration.

2.2 Multilateral index methods

Multilateral index methods have their genesis in comparisons of price levels across countries or regions. These methods satisfy the transitivity, which is a desirable property for spatial comparisons due to the fact that the results are independent of the choice of base country (region). Commonly known methods are the GEKS method (also known as the EKS method – see Gini (1931), Eltetö and Köves (1964), Szulc (1964), the Geary-Khamis (GK) method (Geary, 1958; Khamis, 1972), the CCDI method (Caves, Christensen and Diewert, 1982; Inklaar and Diewert, 2016) or the real time index method (Chessa, 2015)). In this paper, we consider most of these methods but the problem of the best choice of the multilateral formula seems to be still open.

2.2.1 The quality adjusted unit value index and the Geary-Khamis (GK) method

The term "Quality adjusted unit value method" (shortened to the "QU method") was introduced by Chessa (see, for instance, Chessa, 2015, 2016). The QU method is a family of unit value based index methods with the above-mentioned Geary-Khamis (GK) method as a special case. According to the QU method, the price index $P_{QU}^{0,t}$ which compares the period t with the base period 0 is defined as follows:

$$P_{QU}^{0,t} = \frac{\sum_{i \in G_t} p_i^t q_i^t / \sum_{i \in G_0} p_i^0 q_i^0}{\sum_{i \in G_t} v_i q_i^t / \sum_{i \in G_0} v_i q_i^0}, \tag{6}$$

where the numerator in (6) is the measure of the turnover (expenditure) change between the two considered months and the denominator in (6) is a weighted quantity index. Note that both the turnover index and the weighted quantity index are transitive, and thus the price index $P_{QU}^{0,t}$ is also transitive (Chessa et al., 2017). Note also that the quantity weights v_i are the only unknown factors in Formula (6) and these factors convert sold quantities q_i^0 and q_i^t into "common units" $v_i q_i^0$ and $v_i q_i^t$. Prices of products, p_i^0 and p_i^t, are converted into "quality adjusted prices" p_i^0 / v_i and p_i^t / v_i. If the considered consumption segment is homogeneous, then product quantities can be summed (factors v_i are equal for all products) and the index $P_{QU}^{0,t}$ simplifies to the unit value index (the nominator of (6)). If the above-mentioned consumption segment is not homogeneous, then the unit value index must be adjusted. Note also that the formula $P_{QU}^{0,t}$ defines a family of price indices. In fact, limiting considerations to products sold in both moments 0 and t, and setting v_i equal to the product prices in the current period t, the Formula (6) leads to the Laspeyres index:

$$P_{La}^{0,t} = \frac{\sum_{i \in G_{0,t}} p_i^t q_i^0}{\sum_{i \in G_{0,t}} p_i^0 q_i^0}. \tag{7}$$

Similarly, if we consider the group of products $G_{0,t}$ and if the quantity weights v_i are set equal to the prices in the base period (month) 0, then the formula $P_{QU}^{0,t}$ simplifies to the Paasche price index, i.e.

$$P_{Pa}^{0,t} = \frac{\sum_{i \in G_{0,t}} p_i^t q_i^t}{\sum_{i \in G_{0,t}} p_i^0 q_i^t}. \tag{8}$$

In other words, different choices of factors v_i lead to different prices index formulas. In the GK method, the weights v_i are defined as follows:

$$v_i = \sum_{z=0}^{T} \varphi_{i,GK}^z \frac{p_i^z}{P_{QU}^{0,z}}, \tag{9}$$

where:

$$\varphi_{i,GK}^z = \frac{q_i^z}{\sum_{\tau=0}^{T} q_i^\tau}, \tag{10}$$

and $[0, T]$ is the entire time interval of the product observations (typically $T = 12$, see Diewert and Fox, 2017). Please note that Formulas (6), (9) and (10) lead to a set of equations which should be solved simultaneously. The above-mentioned solution can be found iteratively (Maddison and Rao, 1996; Chessa, 2016) or as the solution to an eigenvalue problem (Diewert, 1999). An interesting alternative method for obtaining this solution can be also found in Diewert and Fox (2017).

2.2.2 The augmented Lehr index

The Lehr method is similar to the Geary-Khamis method (see Section 2.2.1, Formula (6) with weights defined in (9)) but it does not use the complex iterative method. The quality adjusted factors v_i are defined here as follows:

$$v_i = \frac{p_i^0 q_i^0 + p_i^T q_i^T}{q_i^0 + q_i^T}. \tag{11}$$

The immediate conclusion from (11) is that the Lehr index uses only data from months 0 and T, and in fact this is a bilateral index. Nevertheless, we can change the formula of the quality adjustment factors, and thus, similarly to multilateral methods, we take into considerations all available information from the interval, i.e. (see Loon and Roels, 2018):

$$v_i = \frac{\sum_{\tau=0}^{T} p_i^\tau q_i^\tau}{\sum_{\tau=0}^{T} q_i^\tau}. \tag{12}$$

In the next part of the paper, the augmented Lehr index, i.e. the index constructed as in (6) with quantity weights defined in (12), will be denoted by $P_{AL}^{0,t}$ and the above-mentioned factors will be signified by v_i^{AL}. In other words, the considered augmented Lehr index can be written as follows:

$$P_{AL}^{0,t} = \frac{\sum_{i \in G_t} p_i^t q_i^t / \sum_{i \in G_0} p_i^0 q_i^0}{\sum_{i \in G_t} v_i^{AL} q_i^t / \sum_{i \in G_0} v_i^{AL} q_i^0} \tag{13}$$

2.2.3 The real time index

Let us note that price imputations are not needed when prices from each month of the current year are included in weights v_i. Taking typically value $T = 12$, Chessa (2015) suggests defining these weights by including product prices and quantities from each month of the current year and the base month December of the previous year (there are 13 months together). However, as the same author admits, in practice, we can use prices and quantities of all 13 months only in the final month of the year, and thus some updating method is needed for v_i calculations each month. Although there are several methods for updating quantity weights (see for instance Krsinich, 2014), we focus on an interesting and quite easy for implementation method proposed by Chessa (2015). He suggests the following procedure of calculating *the real time index*: (1) For the current year, we use a time window with December of the previous year as the fixed base month and the window is enlarged each month with the current month; (2) The price index of the current month t is calculated by using the updated quantity weights according to a special algorithm. In particular, this algorithm needs some initial values of price indices $P_{QU}^{0,\tau} : 0 \le \tau \le t$ and it repeats updating weights $v_i = \sum_{z=0}^{t} \varphi_{i,GK}^z \frac{p_i^z}{P_{QU}^{0,z}}$ and next updating values of price indices $P_{QU}^{0,\tau} : 0 \le \tau \le t$ (according to (6)) until the difference between indices from the last two iterations is small enough. Chessa (2015) recommends a method for calculating initial indices. Moreover, he sets the stop criterion at 0.001 and assumes the maximum absolute difference between the price index vectors as a distance measure. Nevertheless, in our study, we set the stop criterion at 0.0001 and we use the Euclidean distance for comparisons of two successive iterations. Steps (1) and (2) are repeated until December of the current year and after that the base month is shifted to December of the current year. In this way, the whole procedure may be repeated in the subsequent year. For more details, see also Chessa (2016).

2.2.4 The GEKS method

Let us consider a time interval $[0,T]$ of observations of prices and quantities which will be used for the GEKS index construction. The GEKS price index between months 0 and t is an unweighted geometric mean of $T+1$ ratios of bilateral price indices $P^{\tau,t}$ and $P^{\tau,0}$ which are based on the same price index formula. The bilateral price index formula should satisfy the time reversal test, i.e. it should satisfy the condition $P^{a,b} \cdot P^{b,a} = 1$. Typically, the GEKS method uses the superlative Fisher price index and in such case the GEKS formula can be written as follows:

$$P_{GEKS}^{0,t} = \prod_{\tau=0}^{T} \left(\frac{P_F^{\tau,t}}{P_F^{\tau,0}}\right)^{\frac{1}{T+1}}. \tag{14}$$

The GEKS formula based on the Jevons price index is also considered in this paper, i.e.

$$P_{JGEKS}^{0,t} = \prod_{\tau=0}^{T} \left(\frac{P_J^{\tau,t}}{P_J^{\tau,0}}\right)^{\frac{1}{T+1}}. \tag{15}$$

2.2.5 The CCDI method

The GEKS method for making international index number comparisons between countries comes from Gini (1931) but it should be mentioned that it was derived in a different manner by Eltetö and Köves (1964) and Szulc (1964). Feenstra, Ma and Rao (2009), and also De Haan and var der Grient (2011) suggested that the Törnqvist price index formula (see (3)) could be used instead of the Fisher price index in the Gini methodology. Caves, Christensen and Diewert (1982) used the GEKS idea with the Törnqvist index as a base in the context of making quantity comparisons across production units (the CCD method) and Inklaar and Diewert (2016) extended the CCD methodology to making price comparisons across production units. Thus, in the paper of Diewert and Fox (2017), the multilateral price comparison method that uses the GEKS method based on the Törnqvist price index is called the CCDI method. The corresponding CCDI price index can be expressed as follows:

$$P_{CCDI}^{0,t} = \prod_{\tau=0}^{T} \left(\frac{P_T^{\tau,t}}{P_T^{\tau,0}}\right)^{\frac{1}{T+1}}. \tag{16}$$

2.2.6 Other methods

In the literature, we can find some other multilateral index methods which are not considered in this paper. The Country-Product Dummy (CPD) method proposed by Summers (1973) has been adapted for spatial price comparisons to the time domain and now it is known as the Time Product Dummy (TPD) method (de Haan and Krsinich, 2014). The multilateral hedonic method is closely related to the TPD method, i.e. its model parameters (known as "item fixed effects") are not estimated for items (as in the TPD method) but they are estimated for the characteristics of items (attributes). Both the TPD method and the above-mentioned hedonic method do not simplify to a unit value index when all products are homogeneous and they are flawed with regard to their use of turnover in constructing weights (Chessa, 2015). Some other methods can be encountered in the paper of Haan et al. (2016), for instance, the so-called "Cycle Method" (see also Willenborg, 2010, 2017; Willenborg and Van der Loo, 2016).

2.3 Alternative weighting schemes in the QU method

In the classical form, the GK method uses quantity shares as weight in the construction of v_i. In the litera-ture, we can find at least two other weighting schemes in quantity weights for the GK price index. The first variant was proposed by Hill (2000) and it assumes that deflated prices, i.e. $p_i^z / P_{QU}^{0,z}$, are weighted by the ratio of the turnover share of the i-th product in the month z (denoted here by s_i^z) and the sum of turnover shares of the same product over different months. In the paper of Chessa (2016), this variant is referred to as the "QU-TS" method but we use here the shortened notation "TS", i.e. we denote the above-mentioned weights for deflated prices as $\varphi_{i,TS}^z$. In other words, in the TS method, weights $\varphi_{i,GK}^z$ are replaced by weights calculated as follows:

$$\varphi_{i,TS}^z = \frac{s_i^z}{\sum_{\tau=0}^{T} s_i^\tau}, \tag{17}$$

and the final quantity weights are computed as follows:

$$v_i = \sum_{z=0}^{T} \varphi_{i,TS}^z \frac{p_i^z}{P_{QU}^{0,z}}. \tag{18}$$

The other weighting scheme assumes that deflated prices in months with sales receive equal weight, and thus it is denoted here by the EW method (in Chessa (2016), this method is referred to as the "QU-EW" method). In other words, in the considered weighting scheme, we use the following weights for deflated prices:

$$\varphi_{i,EW}^z = \frac{\delta_i^z}{\sum_{\tau=0}^{T} \delta_i^\tau}, \tag{19}$$

where $\delta_i^z = 1$ if $q_i^z > 0$ and $\delta_i^z = 0$ otherwise. Analogically to (17), in the EW method, the final quantity weights can be written as:

$$v_i = \sum_{z=0}^{T} \varphi_{i,EW}^z \frac{p_i^z}{P_{QU}^{0,z}}. \tag{20}$$

In the next part of the paper, we will use different notations for quantity weights defined in (9), (18) and (20), i.e. these weights, connected with the GK, TS and EW methods, will be signified by v_i^{GK}, v_i^{TS} and v_i^{EW} respectively. Similarly, the corresponding multilateral indices, which compare the time moment t with the time moment 0, will be denoted by $P_{GK}^{0,t}$, $P_{TS}^{0,t}$ and $P_{EW}^{0,t}$ respectively. In the paper we also suggest considering a different system of weights based on observed and available expenditures, namely:

$$\varphi_{i,EX}^z = \frac{p_i^z q_i^z}{\sum_{\tau=0}^{T} p_i^\tau q_i^\tau}, \tag{21}$$

which allows us to calculate the final quantity weights in the QU method as follows:

$$v_i = \sum_{z=0}^{T} \varphi_{i,EX}^z \frac{p_i^z}{P_{QU}^{0,z}}. \tag{22}$$

We will denote these quantity weights by v_i^{EX} and the corresponding QU index, i.e. the index defined in (6) but using weights v_i^{EX} instead of weights v_i, by $P_{EX}^{0,t}$.

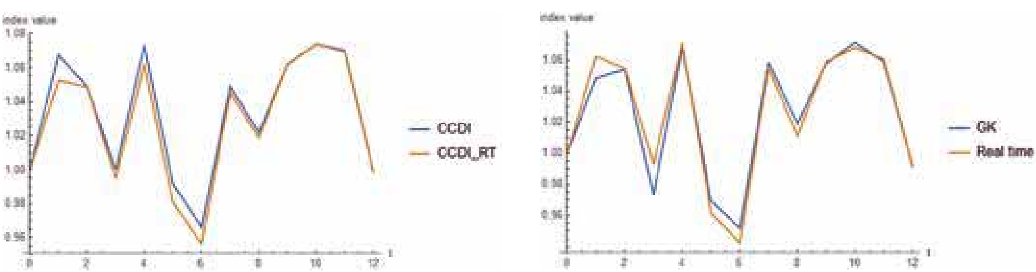
3 EMPIRICAL STUDY

Poland is at the beginning of the way to the regular and official use of scanner data in the CPI measurement. Statistics Poland has started to cooperate with three supermarkets but they do not provide scanner data in a regular way. Moreover, there is no IT system for combining and analysing different data sources from different retailers (supermarkets) written in different file formats. Nevertheless, some experiments on real scanner data sets are being done by using the R package and Mathematica software. In the following empirical study, we consider data sets come from one supermarket chain and they concern the following group of products: plain flour (COICOP group: 011121), milk 3.2% (COICOP group: 011411) and rice (COICOP group: 011111). In this case, we have a 13-month time series (Dec. 2014–Dec. 2015).

Figure 7 presents a comparison of two selected multilateral indices calculated over the whole period of 13 months (i.e. the **CCDI** and **GK** indices when a full window is available) with the corresponding indices calculated over the "currently" available window (i.e. for the current time moment t, the available time window is $[0,t]$ – see the **CCDI_RT** and the **real time** indices). Figure 8 presents a comparison of the **GEKS** index with the **CCDI** and **JGEKS** indices calculated over the whole period of 13 months. Figure 9 presents all considered multilateral indices together with the **chained Jevons** index calculated for the fully available time window. All calculations were done in the Mathematica 11 software.

Figure 7 Comparison of selected multilateral indices (CCDI, GK) for fully and "currently" available time windows (calculated for plain flour, milk and rice)

a) plain flour



b) milk

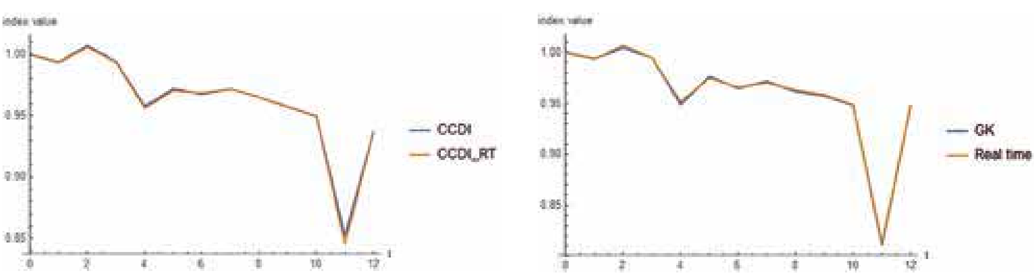
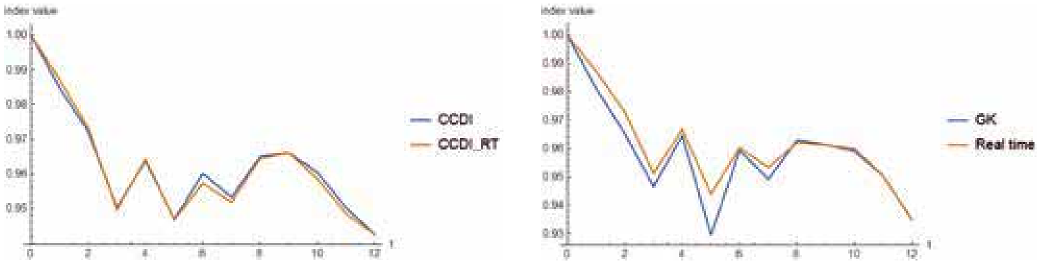


Figure 7 (continuation)

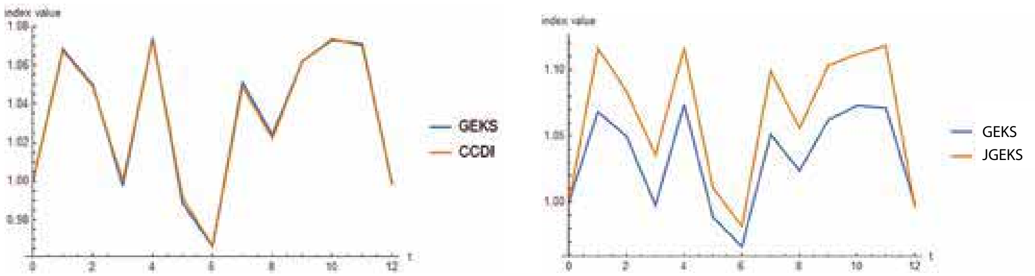
c) rice



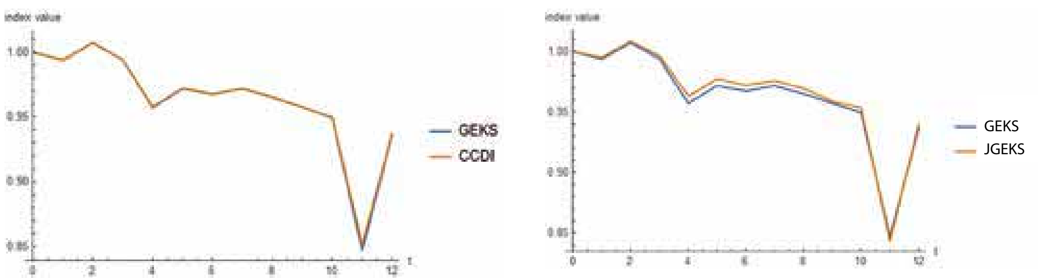
Source: Own calculations based on scanner data from one retailer chain

Figure 8 Comparison of the GEKS index with the CCDI and JGEKS indices calculated over the whole period of 13 months for plain flour, milk and rice

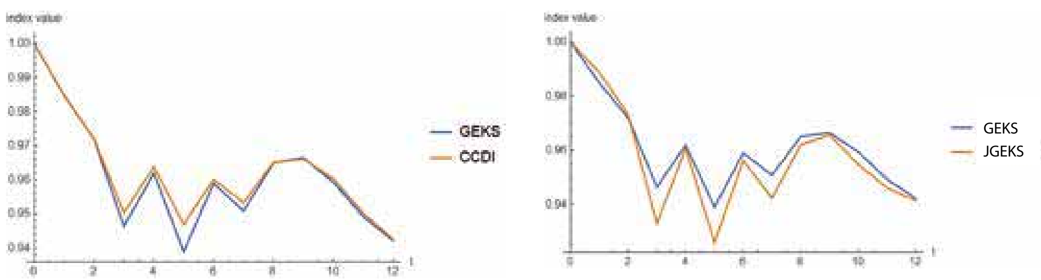
a) plain flour



b) milk



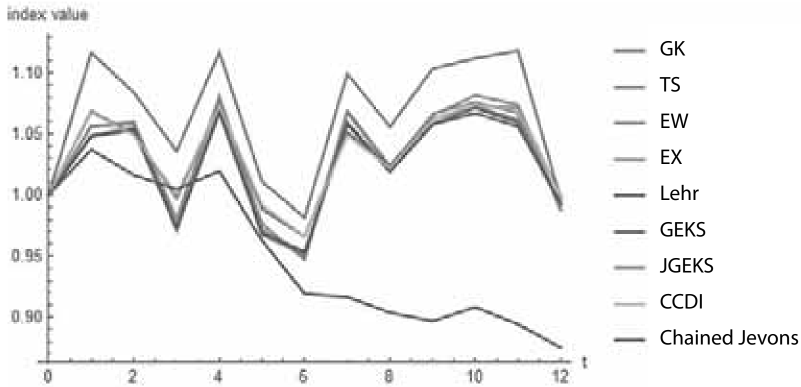
c) rice



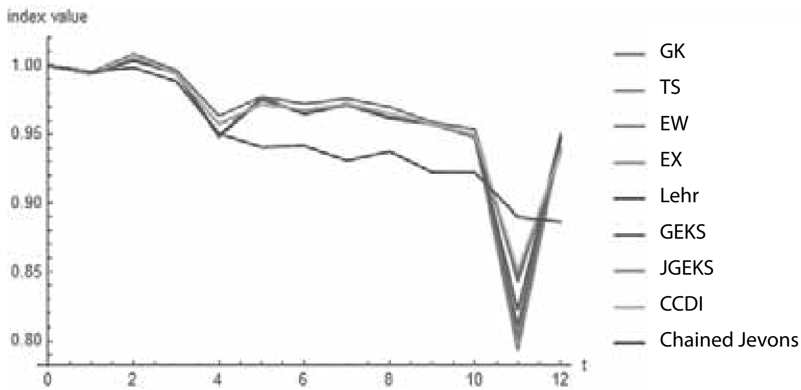
Source: Own calculations based on scanner data from one retailer chain

Figure 9 All considered multilateral indices together with the chained Jevons index calculated over the whole period of 13 months for plain flour, milk and rice

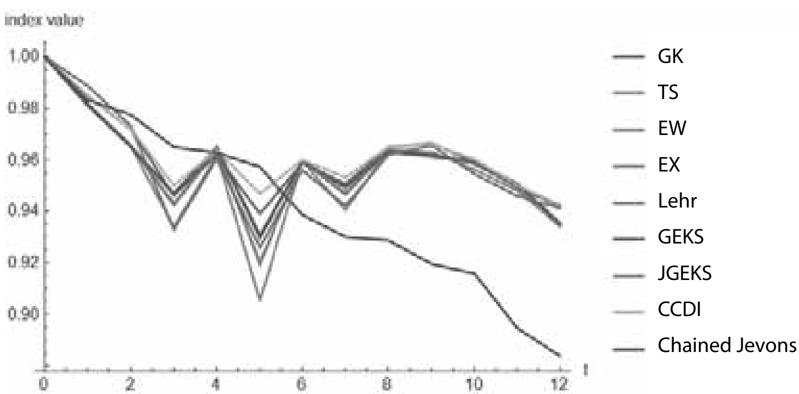
a) plain flour



b) milk



c) rice



Note: For coloured figure see the online version of Statistika journal No. 1/2020.
Source: Own calculations based on scanner data from one retailer chain

CONCLUSIONS

Our empirical study provides the following conclusions: (a) when we have no historical data from supermarkets and we start using scanner data sets, then the application of multilateral indices for the "currently" available time window (from the beginning of cooperation with supermarkets till the current month) is justified since differences between selected indices (CCDI, GK) for the fully and "currently" available time window are not too big, i.e. these differences are decreasing functions of time and, as a rule, after 6–8 months they are negligible (see Figure 7); (b) In practice, there are no substantial differences between the GEKS and CCDI indices and it is not surprising since superlative indices (Fisher, Törnqvist) approximate each other (Diewert, 1976). Nevertheless, the differences between the GEKS and JGEKS indices are crucial and, in our opinion, it confirms that the movements of quantities may not be (rationally) correlated with price movements (see Figure 8); (c) Differences between multilateral indices and the chained Jevons index may be very big (see Figure 9 for plain flour or rice), and as a rule they are. Thus, switching the chained Jevons index to one of multilateral indices does matter in the CPI measurement based on scanner data sets; (d) The choice of the weighting schemes in the QU method does matter – differences in results may be crucial (in our study time moments for which the differences between the TS, EW and EX indices exceeded 3 percentage points were observed).

ACKNOWLEDGEMENTS

This publication is financed by the National Science Centre in Poland (Grant No. 2017/25/B/HS4/00387).

References

ABS. *Making Greater Use of Transactions Data to Compile the Consumer Price Index*. Information Paper 6401.0.60.003, Australian Bureau of Statistics, Canberra, 29 November 2016.

BALK, M. B. Price Indexes for Elementary Aggregates: The Sampling Approach. *Journal of Official Statistics*, 1995, 21(4), pp. 675–699.

CARLI, G. Del valore e della proporzione de'metalli monetati. *Scrittori Classici Italiani di Economia Politica*, 1804, 13, pp. 297–336.

CAVES, D. W., CHRISTENSEN, L. R., DIEWERT, W. E. Multilateral comparisons of output, input, and productivity using superlative index numbers. *Economic Journal*, 1982, 92, pp. 73–86.

CHESSA, A. G. *Towards a generic price index method for scanner data in the Dutch CPI*. Room document for Ottawa Group Meeting, Urayasu City, Japan, 20–22 May 2015.

CHESSA, A. G. A New Methodology for Processing Scanner Data in the Dutch CPI. *Eurona*, 2016, 1, pp. 49–69.

CHESSA, A. G. AND GRIFFIOEN, R. *Comparing Scanner Data and Web Scraped Data for Consumer Price Indices*. Report, Statistics Netherlands, 2016.

CHESSA, A. G. *Comparisons of QU-GK Indices for Different Lengths of the Time Window and Updating Methods*. Paper prepared for the second meeting on multilateral methods organised by Eurostat, Luxembourg, Statistics Netherlands, 14–15 March 2017.

CHESSA, A. G., VERBURG, J., WILLENBORG, L. *A comparison of price index methods for scanner data*. Paper presented at the 15th Meeting of the Ottawa Group on Price Indices, Eltville am Rhein, Germany, 10–12 May 2017.

Consumer Price Index Manual. Theory and practice. IMF, OECD, UNECE, Eurostat, The World Bank, International Labour Office (ILO), Geneva, 2004.

DIEWERT, W. E. Exact and superlative index numbers. *Journal of Econometrics*, 1976, 4, pp. 114–145.

DIEWERT, W. E. *Consumer price statistics in the UK*. Office for National Statistics, Newport, 2012.

DIEWERT, W. E. AND FOX, K. J. *Substitution Bias in Multilateral Methods for CPI Construction using Scanner Data*. Discussion paper 17–02, Vancouver, Canada: Vancouver School of Economics, The University of British Columbia, 2017.

DE HAAN, J. The re-design of the Dutch CPI. *Statistical Journal of the United Nations Economic Commission for Europe*, 2006, 23, pp. 101–118.

DE HAAN, J. AND VAN DER GRIENT, H. A. Eliminating chain drift in price indexes based on scanner data. *Journal of Econometrics*, 2011, 161, pp. 36–46.

DE HAAN, J. AND KRSINICH, F. *Time Dummy Hedonic and Quality-Adjusted Unit Value Indices: Do They Really Differ?* Paper presented at the Society for Economic Measurement Conference, Chicago, U.S, 18–20 August 2014.

DE HAAN J. *A Framework for Large Scale Use of Scanner Data in the Dutch CPI*. Paper presented at the 14th Ottawa Group meeting, Tokyo, Japan, 2015.
DE HAAN, J., WILLENBORG, L., CHESSA, A. G. *An Overview of Price Index Methods for Scanner Data*. Paper presented at the Meeting of the Group of Experts on Consumer Price Indices, Geneva, Switzerland, 2–4 May 2016.
EICHHORN W. AND VOELLER J. *Theory of the Price Index. Fisher's Test Approach and Generalizations*. Berlin, Heidelberg, New York: Springer-Verlag, 1976.
ELTETÖ, Ö. AND KÖVES, P. On a Problem of Index Number Computation Relating to International Comparisons (in Hungarian). *Statisztikai Szemle*, 1964, 42, pp. 507–518.
EUROSTAT. Practical Guide for Processing Supermarket Scanner Data [online]. In: *Harmonised Index of Consumer Prices*, 2017. <https://ec.europa.eu/eurostat/web/hicp/overview>.
FEENSTRA, R. C., MA, H., PRASADA RAO, D. S. Consistent comparisons of real incomes across time and space. *Macroeconomic Dynamics*, 2009, 13(S2), pp.169–193.
FISHER, I. *The Making of Index Numbers*. Boston: Houghton Mifflin, 1922.
GEARY, R. G. A Note on Comparisons of Exchange Rates and Purchasing Power between Countries. *Journal of the Royal Statistical Society*, Series A, 1958, 121, pp. 97–99.
GINI, C. *On the Circular Test of Index Numbers*. Metron 9:9, 1931, pp. 3–24.
GRIFFIOEN, A. R. AND TEN BOSCH, O. *On the Use of Internet Data for the Dutch CPI*. Paper presented at the UNECE-ILO Meeting of the Group of Experts on Consumer Price Indices, Geneva, Switzerland, 2–4 May 2016.
HILL, R. J. Measuring substitution bias in international comparisons based on additive purchasing power parity methods. *European Economic Review*, 2000, 44, pp. 145–162.
INKLAAR, R. AND DIEWERT, W. E. Measuring Industry Productivity and Cross-Country Convergence. *Journal of Econometrics*, 2016, 191, pp. 426–433.
IVANCIC, L., DIEWERT, W. E., FOX, K. J. Scanner Data, Time Aggregation and the Construction of Price Indices. *Journal of Econometrics*, 2011, 161(1), pp. 24–35.
JEVONS, W. S. The variation of prices and the value of the currency since 1782. *J. Statist. Soc. Lond.*, 1865, 28, pp. 294–320.
KHAMIS, S. H. *A* New System of Index Numbers for National and International Purposes. *Journal of the Royal Statistical Society*, Series A, 1972, 135, pp. 96–121.
KRSINICH, F. *The FEWS Index: Fixed Effects with a Window Splice – Non-Revisable Quality-Adjusted Price Indices with No Characteristic Information*. Paper presented at the UNECE-ILO Meeting of the Group of Experts on Consumer Price Indices, Geneva, Switzerland, 2–4 May 2016.
LAMBORAY, C. *The Geary Khamis index and the Lehr index: how much do they differ?* Paper presented at the 15th Ottawa Group meeting, Eltville am Rhein, Germany, 10–12 May 2017.
LASPEYRES, E. Die Berechnung einer mittleren Waarenpreissteigerung. *Jahrbücher für Nationalökonomie und Statistik*, 1871, 16, pp. 296–314.
LEVELL, P. Is the Carli index flawed? Assessing the case for new retail price index RPIJ. *J. R. Statist. Soc.*, 2015, A, 178(2), pp. 303–336.
LOON, K. V. AND ROELS, D. *Integrating big data in the Belgian CPI*. Paper presented at the meeting of the group of experts on consumer price indices, Geneva, Switzerland, 8–9 May 2018.
MADDISON, A. AND RAO, D. S. P. *A Generalized Approach to International Comparison of Agricultural Output and Productivity*. Research memorandum GD-27, Groningen Growth and Development Centre, Groningen, The Netherlands, 1996.
MARTINI, M. A General Function of Axiomatic Index Numbers. *Journal of the Italian Statistics Society*, 1992, 1(3), pp. 359–376.
PAASCHE, H. Über die Preisentwicklung der letzten Jahre nach den Hamburger Borsennotirungen. *Jahrbücher für Nationalökonomie und Statistik*, 1874, 12, pp. 168–178.
SUMMERS, R. International Price Comparisons Based Upon Incomplete Data. *Review of Income and Wealth*, 1973, 19, pp. 1–16.
SZULC, B. Indices for Multiregional Comparisons (in Polish). *Przeglad Statystyczny*, 1964, 3, pp. 239–254.
SZULC, B. Linking Price Index Numbers. In: DIEWERT, W. E. AND MONTMARQUETTE, C. eds. *Price Level Measurement*, 1983, pp. 537–566.
TÖRNQVIST, L. The Bank of Finland's Consumption Price Index. *Bank of Finland Monthly Bulletin*, 1936, 10, pp. 1–8.
VAN DER GRIENT, H. A. AND DE HAAN, J. *The use of supermarket scanner data in the Dutch CPI*. Paper presented at the Joint ECE/ILO Workshop on Scanner Data, Geneva, 10 May 2010.
VON AUER, L. Processing scanner data by an augmented GUV index. *Eurostat Review of National Accounts and Macroeconomic Indicators*, 2017, 1, pp. 73–91.
VON AUER, L. *The Nature of Chain Drift*. Paper presented at the 17th Meeting of the Ottawa Group on Price Indices, Rio de Janeiro, Brasil, 8–10 May 2019.
WALSH, C. M. *The Measurement of General Exchange Value*. New York: The MacMillan Company, 1901.
WILLENBORG, L. *Chain Indexes and Path Independence*. Report, Statistics Netherlands, 2010.
WILLENBORG, L. AND VAN DER LOO, M. *Transitivizing Price Index Numbers Using the Cycle Method: Some Empirical Results*. Report, Statistics Netherlands, 2016.
WILLENBORG, L. *Transitivizing Elementary Price Indexes for Internet Data using the Cycle Method*. Discussion Paper, Statistics Netherlands, 2017.

Income and Consumption Inequalities in Palestine: a Regression-Based Decomposition Approach

Mohsen Ayyash [1] | *Universiti Sains Malaysia, Penang, Malaysia*
Siok Kun Sek [2] | *Universiti Sains Malaysia, Penang, Malaysia*
Tareq Sadeq [3] | *Birzeit University, Birzeit, Palestine*

Abstract

The inequality in the households' living standards is commonly measured by either income or consumption. Different household's attributes may affect inequality in these living standards. This study aims to investigate the factors affecting income and consumption, quantifies their proportionate contributions to income and consumption inequalities, and compares them. The data are collected from the Palestinian Household Expenditure and Consumption Survey (PECS) in 2017. To cast light on this issue, the study applies a regression-based decomposition approach to income-generating function. The results suggest that household attributes better explain adjusted consumption inequality than adjusted income inequality, which should be a better measure of living standards. Moreover, the results indicate that the region, education, and employment status are the major factors of adjusted income and consumption inequalities, while the other factor's contributions have been minimal. For policy interventions, multidimensional policies should be formulated to reduce inequality in all dimensions for achieving an overall equal society.

Keywords	***JEL code***
Inequality, regression-based decomposition, income, consumption, Palestine	*C01, C21, D63*

INTRODUCTION

In recent years, economists' interest in inequality, its dimensions, and decomposition has arisen. Inequality has several dimensions; it can be accompanied by inequality of education, skills, health, opportunities, welfare, access to infrastructure, in addition to inequality of income and wealth. In particular, several dimensions may be linked to economic inequality such as earnings, wages, consumption, expenditure,

[1] School of Mathematical Sciences, University Sains Malaysia, 11800 USM Penang, Malaysia. Corresponding author: e-mail: ayash.mohsen@gmail.com, phone: (+60)174330747.

[2] School of Mathematical Sciences, University Sains Malaysia, 11800 USM Penang, Malaysia. E-mail: sksek@usm.my, phone: (+60)4-6535338.

[3] Department of Economics, Birzeit University, Birzeit, West Bank 627, Palestine. E-mail: tsadeq@birzeit.edu.

and income. According to the economic theory, economic inequality is usually described either in terms of variations among households/ individuals in the distribution of income or consumption within a country, between countries and across geographical regions. Cowell (2009) defined inequality as a scalar numerical value that indicates the discrepancies of an individual's income within a certain population.

A huge body of research is interested in measuring households' living standards, to what extent these living standards are equally distributed, and whether its levels are high or low. Therefore, the income either at household's or per adult equivalent level is commonly used as a proxy in measuring its living standards while consumption, alternatively, is long-preferred by economists. The use of consumption may reflect the actual living standards. Meanwhile, the use of income may give us an actual economic power and measures how households differ in their incomes that come from wages, earnings and self-employment, but it may be under-reported especially for households with little resources. Therefore, income and consumption will differ due to the fact that all households should consume, but not all of them earn income, which in turn leads to higher income inequality compared to consumption inequality. Hence, household living standards should be better measured by household's consumption (Goodman and Oldfield, 2010; Brewer and O'Dea, 2012; Meyer and Sullivan, 2013).

The historical records of inequality in Palestine have been based on the consumption data other than income. Palestine exhibits a decrease in consumption inequality in 2017. That is, the value of the Gini coefficient was 40.3% in 2011, which declined to 34.0% in 2017 (PCBS, 2018). Thus, inequality in Palestine is still observed, but remains low and around the world average.

The present study mainly focuses on studying the distributions of income and consumption. More specifically, this study seeks to give answers to the question of what kinds of sources / factors contribute to inequality and examines their shares in the observed inequality in Palestine. Moreover, this study compares the contribution of each factor to income and consumption inequalities. In terms of econometric settings, the present study applies the Fields (2003) regression-based decomposition approach because of its ability to capture the factor's proportional contribution to the total explained inequality, unlike the traditional methods. This study uses Household Expenditure and Consumption Survey (PECS) for the year 2017 provided by the Palestinian Central Bureau of Statistics (PCBS).

To the best of our knowledge, the study shows the first-ever exploratory results estimated for Palestine by using the regression-based approach. Our main conclusion is that our results confirm that consumption is a better measure of inequality in well-being than income. Moreover, the results from the present study shed light on the role of the region, education, age, gender, employment status, locality, and land ownership in explaining income and consumption inequalities, with the region being among the most important factors that explain both of these inequalities. However, income flows from the urban locality had a decreasing inequality effect. In general, each factor had contributed to different magnitudes to income and consumption inequalities, but each one performs almost similarly for both of inequalities.

To conclude, the findings might provide a shred of strong evidence for government and policymakers to formulate appropriate policies towards an overall improvement of well-being. Such required policies focus on diminishing the regional differences among the West Bank and the Gaza Strip in addition to the redistribution of economic resources among the population, particularly to those with lower incomes, which will lead to higher returns if they are investing in human capital.

The summary of the review of the literature is presented in Section 1. The data and the overview of income and consumption distributions are described in Section 2. The description of the methodology used is shown in Section 3. The regression and decomposition results are interpreted in Section 4. The discussions and conclusion are derived in the last section.

1 LITERAUTRE REVIEW

Recently, different methods have been applied to the decomposition of inequality, which has been largely debated in the literature depending on the raised research question. Heshmati (2004) reviewed most of these methods. Methods to decompose inequality are divided into descriptive and quantitative decomposition methods. The descriptive methods include the decomposition by factor component or sources of income, which allows for measuring the contributions of household's income factors relative to the observed income inequality (Fei et al., 1978; Pyatt et al., 1980; and Shorrocks, 1982), by disjoint population groups as well, which allows for decomposition of income inequality into between- and within-population group components (Pyatt, 1976; Shorrocks, 1984; Cowell and Jenkins, 1995). Such methods answered the questions about what income sources or population groups contribute to inequality. However, the contributions of the household's attribute to income inequality could not be measured and detected using such methods.

On the other hand, the quantitative analysis methods involved regression-based decomposition framework. It was firstly developed by Mincer (1958 and 1970), Becker (1964), Blinder (1973) and Oaxaca (1973), which concerns with estimating the differences in the means of income, where the decomposition had relied on the human capital variables in addition to some other controls. Morduch and Sicular (2002) also implemented a general method to regression-based decomposition. However, the contributions of each factor may differ with the selected inequality index. Wan (2002) relied on Shorrocks (1999) decomposition rule to solve the pitfalls of regression-based inequality decomposition in which there are no restrictions on the regression models and its applicability to be applied to any inequality measure. He showed that the constant and residual terms problems are caused by the methods used in the decomposition of inequality, not by the used index or indices.

However, Fields (2003) proposed a different method, which is applicable to inequality levels or changes and is able to decompose any inequality index. He used a standard semi-logarithmic regression model of income in order to obtain the contributions of different indicators to the changes and the levels of inequality. His approach overcomes several advantages, including its capability to add various predictors in the regression model. Moreover, it considers nonlinear effects and controls for endogeneity. The standard errors of source contributions were also computed. Meanwhile, Bigotta et al. (2015) used Shorrocks (1982) weak consistency assumption to show how to find the shares of Atkinson's inequality index. They revised the Fields (2003) decomposition approach to measure income inequality in terms of Atkinson's index and provided further theoretical results on the contribution of each factor in the regression model.

A number of empirical studies have applied the regression-based decomposition approaches in different countries to decompose inequality (Wan, 2004; Gunatilaka and Chotikapanich, 2009; Naschold, 2009; Manna and Regoli, 2012; Brewer and Wren-Lewis, 2016; Rani et al., 2017; Limanli, 2017; Tripathi, 2018). The finding from these studies revealed that gender, human capital, household size, geographical region, work status, and land ownership are considered to be the most common contributing factors to inequality.

The studies that compare income and consumption are broad. According to Friedman (1957), the household's welfare can be measured using consumption expenditure and may provide more accurate results than income, especially for households with insufficient resources (Blundell and Preston, 1998; Meyer and Sullivan, 2003 and 2013; Krueger and Perri, 2006). A study by Brewer and O'Dea (2012) has shown that if an imputed rent of owned houses is added to the measure of household resources, the average annual growth rates in standards of living will be changed considerably, even after adjusting the price deflator. Moreover, Meyer and Sullivan (2013) concluded that inequality in income was greater than consumption inequality, particularly in the distribution's tails. Despite the aforementioned studies had provided a useful comparison among inequalities in income and consumption, these studies were based on a descriptive analysis. One contribution of the present study is that it is the first study that compares income and consumption inequalities using a regression-based approach.

In the context of Palestine, the literature has, mostly, been based on the household expenditure and consumption rather than the household income. Ramadan et al. (2015) showed that expenditure inequality is mainly explained by education and geographical region using the household's monthly expenditure for the years 2007, 2010, and 2011. On the other hand, the Palestinian central bureau of statistics (PCBS) provided a descriptive analysis of inequality in terms of averages, Lorenz curve, Gini index, and decile ratios using consumption data only.

2 DATA AND OVERVEIEW OF INCOME AND CONSUMPTION IN PALESTINE

This section provides more details about the data used in the present study. Moreover, we will provide an overview of income and consumption distributions in Palestine for the year 2017 based on the data in our hands.

2.1 Data

The present study uses the Household Expenditure and Consumption Survey (PECS) collected by the Palestinian central bureau of statistics (PCBS) from October 2016 to September 2017. The target population in this survey comprised of all households and individuals who were normally lived in Palestine during the recording period 2016–2017. The sampling frame comprised of 532 enumeration areas. The design of the sample is a two-stage stratified cluster sample. In the first step, a random sample of 391 enumeration areas. In the second step, a systematic random sample of 12 households is drawn from each enumeration area selected in the first step.

The data contain information about household's heads monthly income, consumption, and expenditure as well as human capital, gender and some non-human capital such as household size, locality type, geographical region, dwelling, and employment status. The data consist of 3 739 household heads and are weighted by the PCBS sampling weights. Household head consumption data were adjusted by purchasing power to take into consideration the spatial and temporal variations in living costs that might arise when the price of the same goods varies across different locations. This adjustment was already done by the Palestinian central bureau of statistics (see PCBS, 2018, p. 32 for more details). Furthermore, we adjust household income and consumption for family composition (i.e., household size and age of its members) to take into account economies of scale. The present study applies the old OECD equivalence scale (Oxford scale).

Let $n_{adults,i}$ is the number of adults in household i, $n_{children,i}$ is the number of children in household i. The old OECD (Oxford) equivalence scale can be written as follows:

$$n_i = \left[1 + 0.7\left(n_{adults,i} - 1\right) + 0.5 n_{children,i}\right]. \tag{1}$$

The per adult equivalent monthly income is obtained by dividing the total monthly household income by the equivalence scale, n_i. The per adult equivalent monthly consumption is obtained analogously. In Palestine, the average household size is 5.52 individuals in 2017 (PCBS, 2018).

2.2 Overview of income distribution

The average monthly household income in Palestine is 4 586.60 New Israeli Shekels (NIS), while the per adult equivalent monthly income is 1 326.7 NIS in 2017. Since this study seeks to quantify the contribution of population attributes to the total income inequality, we first look at the average adjusted (per adult equivalent) monthly income by population groups that are expected to determine the distribution of income as shown in Table 1. These characteristics are gender, education, region, locality type, employment status, and land ownership. Table 1 indicates the structure of each attribute as well. It seems that the mean adjusted monthly income of males is higher than that of their females counterparts. In other words, males earn 1 330.14 NIS while females earn 1 295.52 NIS. Moreover, education enhances income, that is, the average adjusted monthly income increases monotonically with education levels as shown in Table 1.

Furthermore, the locality type is considered as another factor that affect the distribution of income in Palestine. That is, the mean adjusted monthly income is greater in rural areas while it is lesser in urban and camp zones as shown in Table 1. This is probably because of higher employment rates of Palestinians in the Israeli labour market where the wages are higher and where low to no schooling is required compared with the local employment. In 2018, about 18.4% of Palestinian labor were employed in Israeli labor market as indicated by the report of Palestinian labor force survey 2018 (PCBS, 2018). Moreover, the political division since 2007 and the Israeli siege on Gaza Strip were considered as the main causes of income inequality according to region (ILO, 2018). That is, the mean adjusted monthly income in the West Bank is 1 667.95 NIS compared to 707.14 NIS in the Gaza Strip as indicated in Table 1.

The average per adult equivalent monthly income for employed workers is 1 418.93 NIS while it was 1 098.73 NIS for not employed workers counterparts as shown in Table 1. The later workers are receiving their income either in terms of transfer or assistances/aid programs by the ministry of social affairs and other private agencies in which they are more likely to be poor (PCBS, 2018). On the other hand, a low percentage of land ownership by Palestinians; about 27.0%, which in turn affects their incomes particularly those depend on agricultural income. This is probably due to the restrictions imposed by the wall built by the Israeli government in 2002 which resulted in loss of around 16.8% of the total area of the West Bank. Therefore, this area become fully controlled by the Israeli military in which they make it inaccessible to Palestinians (Hareuveni, 2012). In 2017, the average per adult equivalent monthly income of land owners is 1 705.17 NIS compared to 1 184.61 NIS for their counterparts owning no land as shown in Table 1.

Table 1 Share of positive answers to job search questions and item-response probabilities

Group	No. of obs.	Mean Adjusted Monthly income (S.E)	Mean Adjusted Monthly consumption (S.E)
Gender			
Male	3 363	1 330.14 (22.77)	1 624.76 (53.06)
Female	376	1 295.52 (80.96)	1 432.29 (17.11)
Education			
No education	178	986.64 (62.95)	1 330.32 (59.57)
Lower secondary	2 163	1 212.53 (27.21)	1 373.24 (20.73)
Secondary	584	1 350.24 (55.81)	1 471.43 (42.42)
Associate diploma	240	1 546.18 (67.37)	1 632.21 (65.50)
University	574	1 746.10 (72.24)	1 688.89 (45.56)
Locality			
Urban	2 732	1 327.18 (26.61)	1 473.10 (20.17)
Rural	652	1 541.75 (53.36)	1 570.85 (32.09)
Camp	355	927.83 (40.49)	1 067.84 (38.51)

Table 1 (continuation)

Group	No. of obs.	Mean Adjusted Monthly income (S.E)	Mean Adjusted Monthly consumption (S.E)
Region			
West Bank	2 411	1 667.95 (29.39)	1 798.89 (20.98)
Gaza Strip	1 328	707.14 (23.55)	821.35 (13.98)
Employment status			
Employed	2 662	1 418.93 (25.18)	1 458.21 (18.05)
Not employed	1 077	1 098.73 (43.72)	1 435.44 (34.89)
Land ownership			
Owned	1 020	1 705.17 (53.77)	1 659.38 (31.06)
Not owned	2 719	1 184.61 (22.01)	1 373.69 (18.96)

Note: Standard errors in parenthesis.
Source: Authors calculations by using PECS data, 2017

The simplest way to measure the income inequality can be represented by ranking the population from the poorest to the richest and show the percentage of income associated with each decile of the population. Table 2 shows the patterns of the household total monthly income distribution patterns. It appears from Table 2 that the 10% richest households' monthly income was 10.8 times of the income earned by the 10% poorest households in 2017.

Table 2 The patterns of adjusted monthly income and consumption distributions in Palestine

Poorest	10%	20%	30%	40%	50%	60%	70%	80%	90%	Ratio*
Per cent[a]	1.96	5.6	10.4	16.5	23.9	32.8	43.5	56.5	78.9	10.8
Per cent[b]	4.00	9.4	15.8	23.4	31.8	41.8	53.3	66.5	82.1	4.5

Note: * Ratio of 10% richest to 10% poorest; a: estimation based on per adult equivalent monthly income (adjusted income); b: estimation based on per adult equivalent monthly consumption (adjusted consumption).
Source: Author's calculations by using PECS, 2017

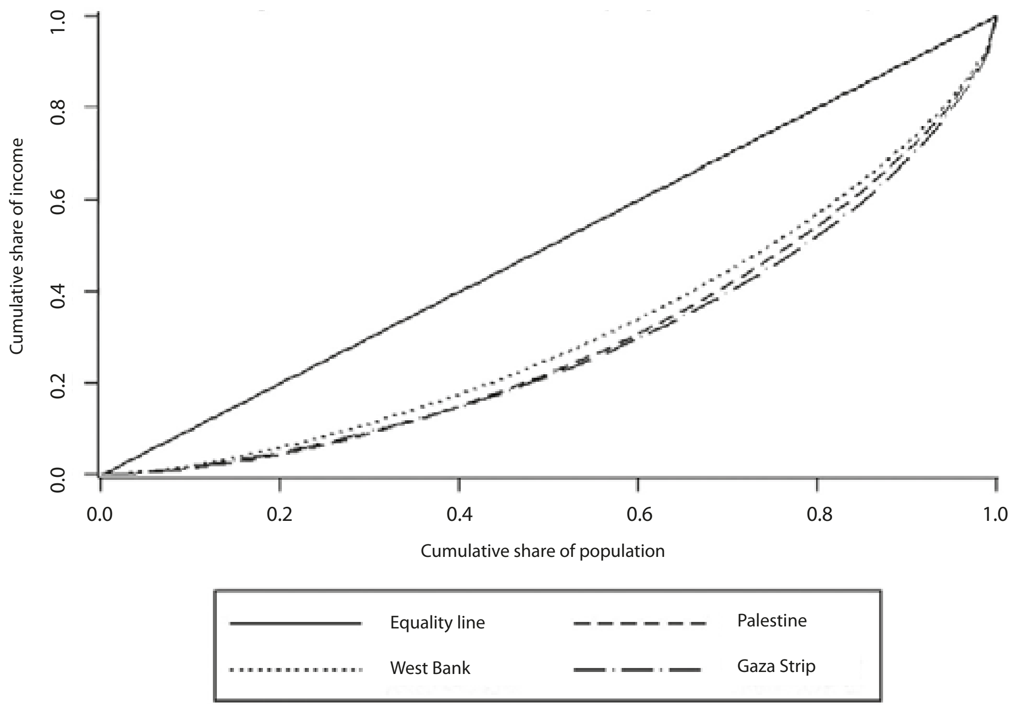
The Gini index of inequality was estimated as 40.98% in Palestine for the year 2017, representing a slight increase in income inequality compared to 40.23% in 2011.[4] The Lorenz curve, shown in Figure 1, is an alternative way to describe income inequality, which is a curve that draws the cumulative percentage of income and the cumulative percentage of ordered population from poorest to richest.

Since Palestine has its own regional characteristics, especially the political division between the West Bank and Gaza Strip and the Israeli blockade on Gaza Strip since 2007, it is worthy to decompose inequality by region, which might provide more insights about regional income inequality in Palestine.

[4] Gini coefficients were calculated using per adult monthly income in both years; 2011 and 2017 based on PECS, 2011 and 2017 data provided by PCBS.

Figure 1 shows the Lorenz curves of both the West Bank and Gaza Strip in addition to Palestine for per adult equivalent monthly income. It appears that income inequality in the Gaza Strip is considerably higher than in the West Bank. Moreover, Theil T inequality index has been estimated for both the West Bank and Gaza Strip since it is additively decomposable. The value of Theil T index in the West Bank is 25.2% while it is 34.3% in Gaza Strip. However, Theil T index for income inequality aggregated in Palestine is 30.6% in 2017 reflecting the extremely large differences among the richest in the West Bank and among the poorest in Gaza Strip (PCBS, 2018).

Figure 1 Lorenz Curve (adjusted income)



Source: Authors construction based on PECS, 2017 data

2.3 Overview of consumption distribution

The consumption usually takes into account the home-produced food and the imputed rent of owned dwelling. A full definition of consumption is provided in the Annex. In 2017, the average monthly household consumption in Palestine was 4 913.3 NIS, while the average monthly consumption per adult equivalent was 1 451.6 NIS. In order to provide a more detailed picture of the distribution of consumption, we will explore it across population attributes as shown in Table 1.

It is evident from Table 1 that the mean per adult equivalent monthly consumption of males is higher than their female counterparts. On the other hand, the inequality in the distribution of consumption is influenced by labour market outcomes such as education, that is, the average monthly consumption per capita rises with education. Ramadan et al. (2015) showed that education is the main determining factor of expenditure inequality for the years 2010 and 2011. Additionally, they found that the composition of the household and the geographical region were the main drivers of expanding the gap in the distribution of expenditure among educated and non-educated household heads.

The average per adult equivalent monthly consumption for those living in camp dwellings is lower compared with their urban and rural counterparts where they show lower average monthly consumption per capita as indicated in Table 1. According to the World Bank (2018) report, Palestine witnessed an increase in the welfare gap between non-camp and camp populations, and the Gini index of inequality in camps increased by 4.0% from 2011 to 2017. Moreover, a divergent regional difference in consumption is evident in Palestine, that is, a high gap exists between the West Bank and Gaza Strip. In 2017, the mean per adult monthly consumption of household living in the West Bank is 1 798.89 NIS compared to 821.35 NIS of their counterparts living in the Gaza Strip. This is pronounced by the high poverty gap at the regional level, which in turn widened the gap in the living standards at the same level. In the Gaza Strip, the poverty had reached 53.0%, while in the West Bank in 2017 it declined to 13.9%. As a result, high concentration of the poor households was in the Gaza Strip, that is, the poverty rate was 71.0% in 2017, compared to 57.0% in 2011 (World Bank, August 2018). Ramadan et al. (2015) also provided evidence that the geographical region was also considered as one of the main determinants of expenditure inequality in 2010 and 2011.

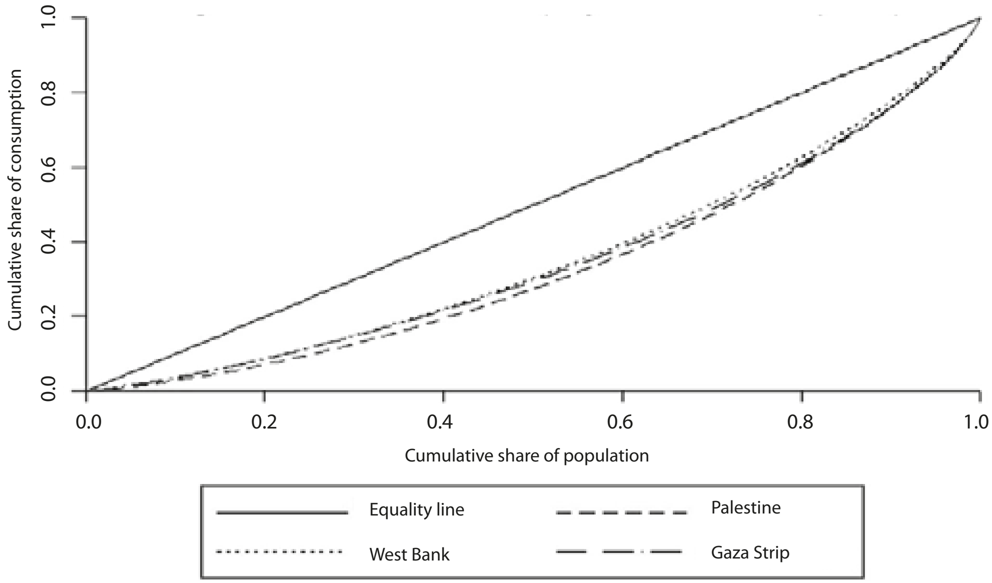
Furthermore, employment status also seems to have effects on the differences in consumption distribution. On average, the per adult equivalent monthly consumption of employed workers is 1.6% higher than that of their not employed counterparts as shown in Table 1. The inequality and poverty had declined over the period 2011 to 2017 in the West Bank due to the changes in labour earnings. In the Gaza Strip, however, the increase in both inequality and poverty rates were due to the reduction in income transfers. Despite the decrease in inequality and poverty by labour earnings, this reduction was not much enough to pay off the decrease in transfers (World Bank, 2018). Furthermore, the differences are also evident by land ownership as shown in Table 1. That is, the mean per adult equivalent monthly consumption of landowners is 1 659.38 NIS while that for non-landowners is 1 373.69 NIS.

In terms of inequality measures, the historical records of inequality in Palestine used consumption as a proxy of income. Palestine experienced a decline in consumption inequality in 2017, where the value of the Gini coefficient fell to 34.0% in 2017 compared to 40.3% in 2011.[5] Moreover, this decline was also exhibited by the decline in the ratio of 10% richest to 10% poorest, which was 4.5% in 2017 compared to 4.8% in 2011 as presented in Table 2 (PCBS, 2018).[6] The regional Gini index varied over the period 2011–2017. It has declined from 39.9% to 31.8%, while Gaza Strip experienced a slight rise from 27% to 28%. However, the inequality in consumption is higher at the country level (World Bank, 2018).[7] Additionally, Figure 2 presents the Lorenz curves of both the West Bank and Gaza Strip in addition to Palestine based on per adult equivalent monthly consumption data.

The inequality in consumption in the Gaza Strip is slightly higher than in the West Bank as illustrated in Figure 2. At the national level, inequality is higher due to the very high consumption gap among the poor in Gaza Strip and the richest in the West Bank. Consumption inequality is lesser than income inequality by about 6.98% points in terms of the Gini index. This is also pronounced if one compares between Lorenz curves exhibited in Figures 1 and 2 in addition to Table 2. Therefore, consumption may provide a better understanding of living standards. In other words, income data lead to larger inequality than consumption data.

[5] Gini coefficients were calculated using per adult equivalent monthly consumption in both years, based on PECS, 2011 and 2017 data provided by PCBS.
[6] Consumption per capita was used in ranking the population and in calculating Gini index.
[7] The corresponding values of Theil T index in Palestine is 17.5%, in the West Bank is 13.5, and in Gaza Strip is 15.2%; calculated by authors based on consumption per capita using PECS, 2017 data.

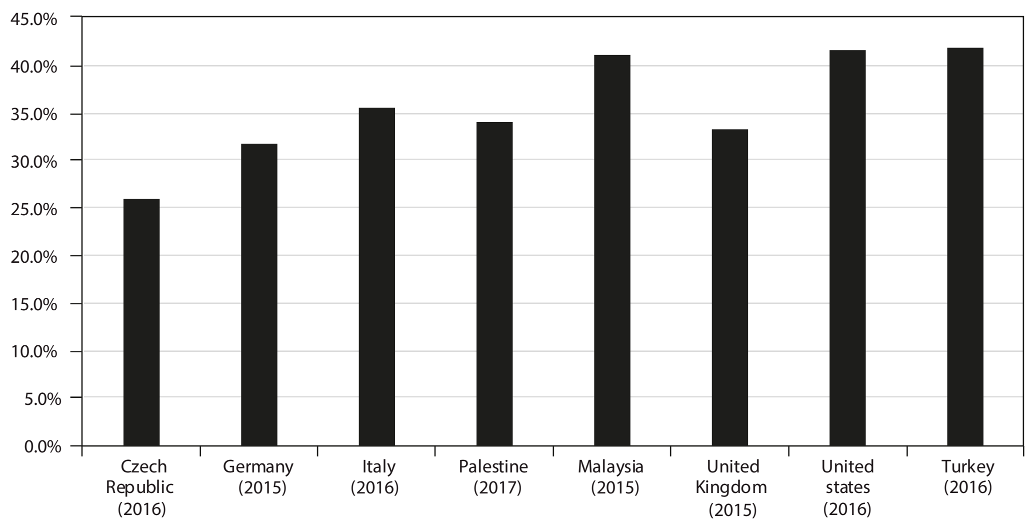
Figure 2 Lorenz Curve (adjusted consumption)



Source: Authors construction based on PECS, 2017 data

Inequality measured by consumption data in Palestine is low by global standards, and it is almost similar to the world average as shown in Figure 3. Estimates were based on the World Bank methodology that is adjusted by purchasing power parity (PPP) for each country.[8]

Figure 3 Gini indices for Palestine and some othe countries



Source: World Bank estimates

8 More information is available at: <*http://iresearch.worldbank.org/PovcalNet/methodology.aspx*>.

3 METHODOLOGY

This paper uses Fields (2003) regression-based decomposition method to examine the contribution of each factor to income and consumption inequalities. This framework firstly applies a linear regression model with different exogenous covariates. That is, income and consumption; measured monthly will be regressed on a set of predictors, which contain some household's attributes such as gender, age, education, etc. Secondly, the proportional contribution of each one of them will be estimated using a specific formula. We briefly present the theoretical framework of this approach.

The starting point of Fields (2003) approach comprises of modelling the function of income, which can be written as:

$$\ln y_i = \beta_0 + \sum_{k=1}^{K} \beta_k x_{i,k} + u_i,\ k = 1,2,\ldots,K, i = 1,2,\ldots,n\ , \tag{2}$$

where, y_i is the outcome variable, which represents the income or consumption and we use the logarithmic transformation to avoid the skewness of the distribution, β_0 is the intercept, $x_{i,k}$ is the k^{th} exogenous factor, βk is the k^{th} coefficient of the k^{th} predictor, u_i is the residual term, and n and k represents the number of observations and the number of predictors, respectively. Formula (2) is a standard linear model, which follows its traditional assumptions (Fields, 2003). For interpretations of the regression results, the estimated regression coefficients from the log-linear models can be exponentiated using the following formula, $(e^{\beta}-1) \cdot 100\%$. For small values of the estimated coefficients, approximately $e^{\beta} \cong \beta$.

Formula (2) will be estimated using the classical approach, i.e., OLS. The results of the estimated Formula (2) are then used to calculate the proportional contribution for each factor k to inequality, which is also known as factor inequality weight denoted by s_k,

$$s_k = \frac{\hat{\beta}_k\, Cov(x_k, \ln y_i)}{Var(\ln y_i)},\ k = 1,2,\ldots,K\, , \tag{3}$$

where $\hat{\beta}_k$ is the estimated coefficient from ordinary least square multiple regression, $Cov(x_k, \ln y_i)$ is the covariance between the outcome variable and the k^{th} predictor, and $Var(\ln y_i)$ is the variance of the outcome variable. The positive sign of s_k exhibits that the k^{th} factor's contribution is an inequality increasing effect while the negative sign of it indicates that it has an inequality decreasing effect. Meanwhile, the value of zero exhibits that the k^{th} factor has no contribution to inequality. The contribution is cumulative in the case of categorical predictors and is estimated by summing the contributions of all respective dummies in the regression equation. In our case, the categorical variables consist of only two categories, which in turn need only one dummy variable. Fields (2003) showed that his decomposition procedure provides a robust method of determining weights to allocate to the several regressors in the linear model given that his six conditions of decomposition are already achieved.[9]

It should be noted that the sum of all factor inequality weights; s_k equal to the coefficient of determination; R-squared. It holds that:

$$\sum_{k=1}^{K} s_k = \frac{\sum_{k=1}^{K} \hat{\beta}_k Cov(x_k, \ln y_i)}{Var(\ln y_i)} = \frac{Var(\ln \hat{y}_i)}{var(\ln y_i)} = R^2\, . \tag{4}$$

[9] The six conditions are listed in the Appendix of Fields (2003).

It remains to show that the proportion of the unexplained inequality refers to the contribution of the residual term, which is denoted by s_u:

$$s_u = 1 - R^2. \tag{5}$$

The advantage of Fields (2003) decomposition procedure is that it can be applied to any inequality index such as the Gini index, variance of log income, and generalized entropy indices. Most importantly, this approach allows for adding any type of independent variables (i.e., categorical and quantitative). Additionally, it measures the contribution of each factor to total inequality by decomposing total consumption or income into components from various factors. However, this procedure is restricted to the log-linear functional form of income data and no contribution of the intercept term to inequality (Wan, 2004).

3.1 Variables

The outcome variables used in the present study are the natural logarithm of both monthly income per adult equivalent and monthly consumption per adult equivalent. The explanatory variables that might influence the distribution of income and consumption comprise of household attributes, which are the gender of household head, education, age, region, locality type, employment status, and land ownership. Table 3 consists of the definitions and descriptive statistics of these variables. The average natural logarithm of monthly income and consumption are 6.82 and 7.07, respectively. Moreover, about 90% of households are headed by males. The average age of the participants is 46.84 years. On average, the household heads have 9.94 years of schooling. The majority of household heads lives in urban areas by 73%. The average percentage of participants from the West Bank is 64.5%, while from Gaza Strip it is 36%. Approximately 71 percent of the households are employed. In terms of land ownership, the average percentage of household heads owned land is 27%, which exhibits that the majority of them are non-landowners.

Table 3 Descriptive statistics of the variables in the study

Continuous variables	Definition	Mean	Std. dev.
Ln (adjusted income)	natural log of total per adult equivalent monthly income in NIS*	6.82	0.90
Ln (adjusted consumption)	natural log of total per adult equivalent monthly consumption in NIS	7.07	0.65
Age	in years	46.84	13.70
Education	completed years of schooling	9.94	4.41
Dichotomous variables	**Definition**	**No. of obs.**	**Per cent[a] %**
Gender	1 for male; 0 for female	3 363	89.9
Urban	1 for urban; 0 for rural and camp	2 732	73.1
Region	1 for West Bank; 0 for Gaza Strip	2 411	64.5
Employment status	1 for employed; 0 for not employed	2 655	71.0
Land ownership	1 for yes, 0 for no	1 010	27.0

Note: * NIS: New Israeli Shekels; a: the percent is referred to the dummy variable with code 1; total number of observations is 3 739.
Source: Author's calculations based on PECS 2017 data

4 RESULTS

4.1 Regression results

In this study, we propose two different models. In model I, the natural logarithm of adjusted monthly income is regressed on the set of predictors mentioned earlier. In model II, the natural logarithm of adjusted monthly consumption is regressed on the same set of predictors. We start the decomposition by estimating both models using the ordinary least square method (OLS). The results of the estimated regression coefficients are presented in Table 4. For reliability, we report estimated regression coefficients with their respective standard errors in parenthesis.

Model I is expected to examine the determinants of income. It is found that model I explained about 37.6% of the total variability in the natural logarithm of monthly income per adult equivalent. On the other hand, the determinants of consumption are examined using model II in which this model explained about 41.6% of the total variability in the natural logarithm of monthly consumption per adult equivalent. However, both models are well fitted and passed all diagnostic tests. Additionally, both models performed well as indicated by their respective values of the adjusted R-square, which are conventional in this type of studies. Our results are in line with what was reported in the literature (Gunatilaka and Chotikapanich, 2009; Manna and Regoli, 2012; Bigotta et al., 2015; Rani et al., 2017; Tripathi, 2018). It seems that all predictors are positive and highly significant in both models except land ownership dummy in model II. In other words, these predictors are considered to be the determinants of income and consumption distributions.

Table 4 Multiple regression results for models I and II

Variable	Model I (Adjusted Income)	Model II (Adjusted Consumption)
	Mean (S.E)	Mean (S.E)
Intercept	4.761*** (0.081)	5.829*** (0.056)
Gender	0.152*** (0.044)	0.181*** (0.031)
Education (years)	0.049*** (0.003)	0.037*** (0.002)
Age (years)	0.014*** (0.001)	0.008*** (0.001)
Urban locality	0.059** (0.027)	0.123*** (0.019)
Region	0.887*** (0.026)	0.798*** (0.018)
Employment status	0.550*** (0.033)	0.092*** (0.023)
Land ownership	0.085*** (0.027)	0.031 (0.019)
R-Squared	0.377	0.418
Adj. R-Squared	0.376	0.416
F-Statistic	322.09	382.12
P-value	0.000	0.000

Note: *** significant at 1%, ** significant at 5%, * significant at 10%.
Source: Author's calculations based on PECS 2017 data

The gender dummy is highly statistically significant in both models. The results show that, on average, per capita equivalent income for males is 16.4%[10] higher than their female counterparts. On the other hand, the average per capita equivalent monthly consumption of males is 19.8% higher than female counterparts. That is, the gender consumption gap is higher than the gender income gap.

In both models, education has highly positive and statistically significant effects, which shows that higher income and consumption are associated with higher levels of education. The average per capita equivalent monthly income increases by 5% as an individual's education increase by one year. Meanwhile, the average per capita equivalent consumption increases by 3.8% associated with one unit increase in schooling. Moreover, age is significant on both income and consumption, but this effect is negligible.

It is also found that the urban dummy is statistically significant in both models. That is, on average, a household resides in an urban locality earn monthly income 6% higher than those living in rural and camp localities counterparts. However, the average per capita equivalent monthly consumption of households living in urban areas is 13% higher than those living in rural and camp localities.

The coefficient of region dummy is strongly statistically significant, which indicates the presence of regional influence on both per capita equivalent income and per capita equivalent consumption distributions. This means that a considerable gap exists among those residing in the West Bank and Gaza Strip. The per capita equivalent income gap is 143%, while per capita equivalent consumption gap is 120% in favour of the West Bank. According to the income gap among the employed and not employed, the results reveal that, on average, per capita equivalent monthly income of employed workers is 73.3% higher than their not employed counterparts. However, the consumption gap is 9.6%, which is lower the income gap due to the fact that households might borrow to pay for their consumption, where these results are highly statistically significant.

The dummy of land ownership is statistically significant in model I. The results show when compared to non-land ownership; i.e., reference category, individuals owned land have positive income differentials. On average, individuals owned land receive per capita equivalent monthly income 9% higher than their counterparts owning no land. However, this effect is not significant on per capita equivalent consumption; model II.

Finally, the above results confirmed with what we have reported on the descriptive statistics and the earlier detailed discussions.

4.2 Income inequality decomposition

This section quantifies the contribution of all statistically significant predictors to the explained inequality in order to decide the most important ones in accounting for income inequality in Palestine in 2017. The estimation results of model I are used to calculate factor inequality weights based on Formula (2). Income inequality decompositions are illustrated in Table 5 based on Fields (2003) approach measured by the variance of log monthly income per adult equivalent. For simplicity, each factor inequality weight is divided by the explained income inequality (i.e., R-squared). The resulted percentages exhibit the proportional contributions of different household characteristics to the total explained inequality. It should be noted that even though all predictors are statistically significant in model I, their proportional contribution to the total explained inequality varies considerably.

The results suggest that the region emerges as the most dominant factor contributing to the total explained income inequality in Palestine. The contribution of the region to income inequality is 52.41%.

[10] This figure is obtained by $(e^{0.152}-1) \cdot 100\% = 16.4\%$. The remaining coefficients are interpreted analogously. Refer to Section 3 for more details.

The contribution of the employment status of the household head is captured by the employment status dummy, which substantially contributes to the explained income inequality in which its share reached 22.97%. Moreover, the contribution of education is 18.17%, which seems to be another major contributing factor to income inequality.

On the other hand, income flows from land ownership dummy contributed about 3.42%, while from gender dummy contributed about 2.23%. The contribution from age has been minimal (i.e., 0.98%). Most importantly, the urban dummy showed negative contributions to total explained inequality and thus has inequality decreasing effects. Finally, unobserved factors (i.e., residuals) contributed by about 62.3%.

Table 5 Proportional contribution of each factor to inequality

Variable	Model I (Adjusted Income)		Model II (Adjusted Consumption)	
	s_k%	% of R^2	s_k%	% of R^2
Gender	0.84	2.23	1.08	2.58
Education	6.85	18.17	7.45	17.82
Age	0.37	0.98	0.62	1.48
Urban locality	–0.07	–0.18	0.78	1.87
Region	19.76	52.41	28.39	67.92
employment status	8.66	22.97	3.26	7.80
Land ownership	1.29	3.42	–	–
Residual	62.30	–	58.42	–
Total	100.00	100.00	100.00	100.00

Note 1: % of R^2 – percentage contribution of the factor from the total explained inequality.
Note 2: The contribution of a non-significant predictor should not be considered and thus its contribution should be added to residuals and replaced by '–'.
Source: Author's calculations based on PECS 2017 data

4.3 Consumption inequality decomposition

We follow the same procedure in the previous section to determine the most important factors affecting the distribution of consumption in Palestine for the year 2017. The estimated regression results from model II is used to calculate factor inequality weights. All factors appear to have contributed positively to consumption inequality except land ownership dummy. The results indicate that the region accounted for the largest contribution to the total explained inequality by about 67.92%. This is followed by education, which accounts for 17.82% to the total explained consumption inequality. A considerable lower shares of employment status to consumption inequality if compared with its shares to income inequality, that is, the contribution of employment status is 7.80%.

However, the remaining variables with a small proportional contribution to consumption inequality are gender (2.58%) and age (1.48%). While urban dummy had positively accounted for consumption inequality, its contribution is relatively small, i.e., 1.87%. Lastly, the contribution of unobserved factors (i.e., residuals) is 58.42%, which is lower than the unexplained part of income inequality.

DISCUSSIONS AND COCLUSIONS

This study has applied the regression-based decomposition method developed by Fields (2003) to examine the extent to which different factors contribute to income and consumption inequalities in Palestine. It compares these inequalities as well. This approach is more preferred to other traditional decompositions because of its ability to include several factors in the decomposition model and its applicability to any inequality measure.

The results of the present study confirm various previous results yielded by the most recent literature (Gunatilaka and Chotikapanich, 2009; Naschold, 2009; Manna and Regoli, 2012; Ramadan et al., 2015; Bigotta et al., 2015; Rani et al., 2017; Limanli, 2017; Tripathi, 2018). The study concluded that the major contributing factors to income and consumption inequalities in Palestine are the region, education, employment status; with the region showing the highest contribution, which is similar to Turkey as shown by Limanli (2017). However, the region has very little contribution to inequality in Italy, India, and Sri Lanka (Gunatilaka and Chotikapanich, 2009; Manna and Regoli, 2012; Bigotta et al., 2015; Rani et al., 2017). The reason behind that region has the highest contribution percentage is due to the Israeli siege imposed on the Gaza Strip, where only workers from the West Bank are permitted to enter the Israeli labour market where the wages are very high compared with the domestic labour market. Additionally, the political division between the West Bank and Gaza Strip since 2007, which troubled income earners in Gaza Strip.

Meanwhile, almost similar are the contributions of education to income and consumption inequalities; i.e., 17.82–18.17% and similar to what was reported in the literature. However, minimal contributions to inequality in both income and consumption had attached to the remaining factors, which is in line with the literature as well. Our results reveal that the role of gender, however, is relatively small in Palestine compared to Italy in which gender accounted for 21.3% to Italian income inequality as shown by Manna and Regoli (2012). One possible explanation of the small contribution of gender to inequality in Palestine might be attributed to the prevailing traditions that most households are headed by males. Additionally, most households headed by females are either widowed or divorced or separated, which is reflected in our study sample in which the proportion of female household heads comprises of about 10.1% and the proportion of widowed and divorced female heads is about 80% of them.

Moreover, the analyses provide a useful comparison between income and consumption inequalities. Consumption inequality is lesser than income inequality in Palestine when adjusted for household size, which confirms that the living standards is better measured by consumption than income, which is reflected in terms of higher model fit, i.e., R-squared is 41.6% in model II while it is 37.6% in model I, and in terms of inequality measures, i.e., the Gini for per adult equivalent consumption is 34.0% while it is 40.98% for per adult equivalent income in 2017. This result is similar to the results found in the U.S and the U.K (Goodman and Oldfield, 2010; Meyer and Sullivan, 2013). It should be noted that all factors have almost the same contributions to inequality in Palestine except employment status. The inequality weight of employment status to consumption inequality is relatively smaller than its weight to income inequality. This is probably due to the fact that most unemployed households receive their income in terms of either transfer or social assistance programmes (PCBS, 2018). However, their consumption does not differ considerably from their employed household head counterparts.

Lastly, the findings from the present study are useful for policy implications. The study provides evidence about the drivers of inequality in living standards, which encourage policymakers to prioritize designing policies such as eliminating the regional differences and redistribution of economic resources to protect living standards and the welfare in Palestine and thus moving towards an overall improvement of well-being and a more equitable society. Enhancing education, particularly for those with lower income would lead to higher returns and thus reducing inequality as well.

References

BECKER, G. S. *Human Capital: a theoretical and empirical analysis*. New York: National Bureau of Economic Research, 1964.

BIGOTTA, M., KRISHNAKUMAR, J., RANI, U. Further results on the regression-based approach to inequality decomposition with evidence from India. *Empirical Economics,* 2015, 48, pp. 1233–1266.

BLINDER, A. S. Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, 1973, 8(4), pp. 436–455.

BLUNDELL, R. AND PRESTON, I. Consumption inequality and income uncertainty. *Quarterly Journal of Economics*, 1998, 113, pp. 603–40.

BREWER, M. AND O'DEA, C. *Measuring living standards with income and consumption: evidence from the UK, Institute for social and economic research*. Working paper No. 5/2012.

BREWER, M. AND WREN-LEWIS, L. Accounting for changes in income inequality: decomposition analyses for the UK, 1978–2008. *Oxford Bulletin of Economics and Statistics*, 2016, 78(3), pp. 289–322.

COWELL, F. A. *Measuring Inequality*. 3rd Ed. Oxford: Oxford University Press, December 2009.

COWELL, F. A. AND JENKINS, S. P. How much inequality can we explain? A methodology and an application to the USA. *Economic Journal*, 1995, 105, pp. 421–430.

FEI, J. C. H., RANIS, G., KUO, S. W. Y. Growth and the family distribution of income by factor components. *Quarterly Journal of Economics*, 1978, 92, pp. 17–53.

FIELDS, G. S. Accounting for income inequality and its change: a new method, with application to the distribution of earnings in the United States. *Research in Labour Economics*, 2003, 22, pp. 1–38.

FRIEDMAN, M. *A Theory of the Consumption Function.* National Bureau of Economic Research, Inc., 1957.

GOODMAN, A. AND OLDFIELD, Z. *Permanent Differences? Income and Expenditure Inequality in the 1990s and 2000s*. The Institute for Fiscal Studies, June, 2004.

GUNATILAKA, R. AND CHOTIKAPANICH, D. Accounting for Sri Lanka's Expenditure Inequality 1980–2002: Regression Based Decomposition Approaches. *Review of Income and Wealth*, 2009, 55(4), pp. 882–906.

HAREUVENI, E. *Arrested Development: The Long Term Impact of Israel's Separation Barrier in the West Bank*. Technical Report, Jerusalem: BeTSELEM – The Israeli Information Centre for Human Rights in the Occupied Territories, 2012.

HESHMATI, A. A Review of Decomposition of Income Inequality. *IZA discussion paper No. 1221*, 2004.

INTERNATIONAL LABOUR ORGANIZATION. *The Occupied Palestinian Territory: An Employment Diagnostic Study*. Regional Office for Arab States, Beirut: ILO, 2018.

KRUEGER, D. AND PERRI, F. Does income inequality lead to consumption inequality? Evidence and theory. *The Review of Economic Studies*, 2006, 73(1), pp. 163–193.

LIMANLI, O. Accounting for Income Inequality in Turkey: Regression Based Decomposition Approach. *Journal of Academic Researches and Studies*, 2017, 9(17), pp. 137–146.

MANNA, R. AND REGOLI, A. Regression-based approaches for the decomposition of income inequality in Italy, 1998–2008. *Rivista di statistica ufficiale*, 2012, 14(1), pp. 5–18.

MEYER, B. D. AND SULLIVAN, J. X. Measuring the Well-Being of the Poor Using Income and Consumption. *Journal of Human Resources*, 2003, 38, pp. 1180–1220.

MEYER, B. D. AND SULLIVAN, J. X. Consumption and Income Inequality and the Great Recession. *American Economic Review: Papers & Proceedings*, 2013, 103(3), pp. 178–183.

MINCER, J. Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 1958, 66, pp. 281–302.

MINCER, J. The Distribution of Labour Incomes: a survey with the special reference to the human capital approach. *Journal of Economic Literature*, 1970, 8(1), pp. 1–26.

MORDUCH, J. AND SICULAR, T. Rethinking inequality decomposition, with evidence from rural China. *Economic Journal*, 2002, 112(476), pp. 93–106.

NASCHOLD, F. Microeconomic Determinants of Income Inequality in Rural Pakistan. *Journal of Development Studies*, 2009, 45(5), pp. 746–768.

OAXACA, R. Male–female wage differentials in urban labour markets. *International Economic Review*, 1973, 14, pp. 693–709.

PALESTINIAN CENTRAL BUREAU OF STATISITICS. *Main Findings of Living Standards in Palestine in 2017 (Expenditure, Consumption and Poverty)*. Ramallah – Palestine: PCBS, 2018.

PYATT, G. On the interpretation and disaggregation of Gini coefficients. *Economic Journal*, 1976, 86, pp. 243–255.

PYATT, G., CHEN, C., FEI, J. The Distribution of Income by Factor Components. *The Quarterly Journal of Economics*, 1980, 95(3), pp. 451–473.

RAMADAN, R., HLASNY, V., INTINI, V. Inequality Decomposition in the Arab Region: Application to Jordan, Egypt, Palestine, Sudan and Tunisia. *IARIW-CAPMAS Special Conference: Experiences and Challenges in Measuring Income, Wealth, Poverty and Inequality in the Middle East and North Africa*, Cairo, Egypt, 2015.

RANI, U., KRISHNAKUMAR, J., BIGOTTA, M. Accounting for income inequality: empirical evidence from India. *Indian Economic Review*, 2017, 52(1–2), pp. 193–229.
SHORROCKS, A. F. *Decomposition Procedures for Distributional Analysis: a Unified Framework Based on the Shapley Value* (unpublished manuscript). Department of Economics, University of Essex, 1999.
SHORROCKS, A. F. Inequality decomposition by factor components. *Econometrica*, 1982, 50(1), pp. 193–211.
SHORROCKS, A. F. Inequality decomposition by population subgroups. *Econometrica*, 1984, 52(6), pp. 1369–1385.
TRIPATHI, S. Source of Inequality in Consumption Expenditure in India: A Regression-Based Inequality Decomposition Analysis. *Journal of Income Distribution,* 2018, pp. 1–34.
WAN, G. Regression-based Inequality Decomposition: Pitfalls and a Solution Procedure. *WIDER Working Paper Series 101*, *World Institute for Development Economic Research (UNU-WIDER)*, 2002.
WAN, G. Accounting for Income Inequality in Rural China: A Regression Based Approach. *Journal of Comparative Economics*, 2004, 32, pp. 348–63.

ANNEX

Here, we provide the definitions of household, household head, household income and household consumption as they defined by the Palestinian central bureau of statistics (PCBS).

HOUSEHOLD

'A household is defined as a group of persons who share the same living accommodation, who pool some, or all, of their income and wealth and who consume certain types of goods and services collectively, mainly housing and food. Households are mainly consumers, but they may also be producers. All economic activity taking place within the production boundary and not performed by an entity maintaining a complete set of accounts is considered to be undertaken in the household sector'.

HOUSEHOLD HEAD

'A person who generally provides the chief source of income for the household unit. He is the adult person, male or female, who is responsible for the organization and care of the household or who is regarded as such by the members of the household'.

INCOME

'Cash or in-kind revenues to an individual or household within a given period of time: could be a week, a month, or a year'.

HOUSEHOLD CONSUMPTION

'It refers to the amount of cash spent on purchase of goods and services for living purposes, and the value of goods and service payments or part of payments received from the employer, and own-produced goods and food, including consumed quantities during the recording period, and imputed rent for owned houses'.

Profiling: a New Way to Increase the Quality of Statistics on Research and Development

Thomas Balcone [1] | *Ministry for Higher Education, Research and Innovation, Paris, France*

Abstract

Currently, statistics on Research and Development (R&D) carried out in the business sector are computed in France on the sole basis of legal units: firstly, a survey is addressed to them to collect the data and then, statistics on R&D are disseminated at legal unit level. Considering the increasing importance of the enterprise group in the French economy, it seems difficult today to go on using only the legal units to calculate business statistics. Indeed, assimilating the legal unit to the enterprise is not relevant anymore for group's affiliates and subsidiaries. Taking into account the European definition of an enterprise will help to disseminate more consistent and relevant R&D statistics on the business sector.

The French business statistical register established by the French National Statistical Institute (INSEE), called SIRUS, contains notably all the legal units and all the enterprises. The main contribution of this register is to make possible the calculation and dissemination of statistics at another level than the legal unit one: the enterprise level.

This article first describes why the data should go on being collected at the legal unit level and not at the enterprise one. Indeed, it seems that such a change in the data collection can be dangerous because it could result in a substantial increase of the response burden. Then, this article presents the process based on SIRUS that leads to the computation of key indicators on R&D at enterprise level. To conclude, it compares these key indicators with the ones calculated at the legal unit level to show the impact of moving to the enterprise level on French R&D statistics.[2]

Keywords

Business statistics, business R&D statistics, statistical unit, data collection, surveys

JEL code

M21, C46, O32, C18, C83

INTRODUCTION

In the context of globalization, Research and development (R&D) is a major issue for firms and countries to stay competitive by bringing notably innovation. The globalization of corporate R&D has become

[1] Ministry for Higher Education, Research and Innovation, 1 rue Descartes, Paris, 75231 Île-de-France, France. E-mail: thomas.balcone@recherche.gouv.fr, thomas.balcone@enseignementsup.gouv.fr.

[2] This contribution was presented at the *European Conference on Quality in Official Statistics* (Q2018), Krakow, Poland, 26–29 June 2018.

a considerable phenomenon in the last few years. Therefore, there is a huge need to collect consistent data on R&D to follow properly the evolution of R&D expenditure, notably in the business sector. This is the goal of the R&D survey, which is based on the Frascati Manual, reference work of the OECD which gives all the countries common method and definitions for the conduct of a R&D survey.

In France, this survey led by the Ministry of Higher Education, Research and Innovation since 1963, is based on questioning legal units. With the globalization and the more and more noticeable presence of big groups in the business sector, a new definition of the enterprise appeared in the European regulation,[3] more based on economy than legality: an enterprise is the smallest combination of legal units that is an organizational unit producing goods and services, enjoying a certain decision-making autonomy, especially for the allocation of its current resources. Therefore, computing R&D data at this enterprise level seems to be more relevant to study R&D in the business sector and could become soon mandatory with the implementation of the future framework regulation for the business statistics FRIBS[4] which is currently under negotiation.

In the first part of this article, we will see how we can get data at the enterprise level from the R&D survey. Then, the article presents a method to compute key indicators on R&D at this new level. To conclude, it compares these key indicators with the ones calculated at the legal unit level to show the impact of moving to the enterprise concept on French R&D statistics.

1 HOW TO GET THE DATA AT THE ENTERPRISE LEVEL?

1.1 First approach: collect the data at the enterprise level

Currently, the data collection unit in the French R&D survey is the legal unit. The most natural idea to compute statistics on R&D at the enterprise level with the aim to disseminate more consistent and relevant statistics, is to choose the enterprise as the new data collection unit. However, such a change of data collection unit presents two major risks.

The first risk is a drop of our response rate which is currently over 90%. The legal unit considered as the decision-making unit of the enterprise is not necessarily the one which conducts the R&D activity, and in some large enterprises several legal units can be active in R&D. This can be an issue if we have only one contact in the decision-making unit and may probably increase the response burden for the enterprises. As a consequence, the response rate for large enterprises with several legal units active in R&D, which constitute a major part of the total internal expenditure on R&D, could drop considerably. Moreover, currently, with the legal unit as data collection unit, it's already hard for some contacts to collect data in all the establishments making up the legal unit.

The second risk is the one of no longer being able to do an analysis at the legal unit level. This kind of analysis is relevant to understand how the enterprises organize their R&D activity. For example, it is interesting to know if this activity is located in a sole legal unit (or in a dedicated R&D centre) or divided between several legal units. In the French R&D survey on the 2015 data, nearly 28% of the enterprises likely to carry out R&D activities[5] have more than one legal unit in the survey population, i.e. the population of the legal units likely to carry out R&D activities. Moreover, the legal unit concept is used in French national accounts and there will be no change in the near future. So, it is essential to keep computing consistent and relevant R&D statistics at the legal unit level.

In conclusion, choosing the enterprise as the new data collection unit does not seem the best way to get the best analysis of the R&D activities carried out by the business sector. So, we keep the legal unit as the data collection unit.

[3] The European regulation 696/93.

[4] Framework Regulation Integrating Business Statistics.

[5] An enterprise is likely to carry out R&D activities if at least one of its legal units is in the survey population.

1.2 Second approach: rebuild the enterprises' profiles from legal units

The setting up of the R&D survey population is atypical: contrary to the other surveys conducted in the business sector in France, this population does not come directly from the French business statistical register, called SIRUS, established by INSEE,[6] which contains notably all the legal units and all the enterprises. According to the recommendations of the Frascati Manual 2015[7] (cf. p. 220–221), the target population of the R&D survey at the legal unit level, noted *U(LU)*, is only all legal units likely to perform R&D. This population is built up every year by using the population of the last survey and several other sources linked to R&D or innovation listing legal units (tax credit claimants, innovation survey (CIS),[8] young innovative firms aids…). As the R&D activity carried out by an enterprise can be considered as all the R&D activities performed by its legal units, the target population at the enterprise level, noted *U(EP)*, is naturally all enterprises for which at least one legal unit is likely to perform R&D, i.e. belonging to the target population *U(LU)*.

Following the same logic, the sample at the enterprise level, noted *S(EP)*, is made by all enterprises for which at least one legal unit belongs to the sample *S(LU)*. To compute data on R&D at the enterprise level from the sample *S(EP)*, it is necessary to know the data for all the legal units belonging to an enterprise in the sample and to the target population *U(LU)*. Unfortunately, the data are not available for all these units because first of all, the sample of the R&D survey *S(LU)* does not come from a cluster sampling with the enterprise as the cluster. Indeed, as for most of the other surveys conducted in the business sector in France, the target population at the legal unit level *U(LU)* is divided into an exhaustive stratum and a non-exhaustive one. The exhaustive stratum is made up of the large[9] legal units and of the ones which appear for the first time in the target population. All the other legal units of the target population form the non-exhaustive stratum. In the 2015 survey, the respondent legal units correspond to 1 435 enterprises in *S(EP)* made up of more than one legal unit in the target population *U(LU)* (cf. Table 1). 1 090 of these enterprises are formed by at least one legal unit which is in the target population *U(LU)* but whose R&D

Table 1 The target population at the legal unit level *U(LU)* and the enterprise level *U(EP)* in the 2015 R&D survey

<table>
<tr><th colspan="2">Population</th><th>Legal unit (LU) level</th><th colspan="4">Enterprise level</th></tr>
<tr><td colspan="2">Target population</td><td>U(LU): 25 962</td><td colspan="4">U(EP): 21 466</td></tr>
<tr><td rowspan="5">Sample population</td><td rowspan="3">Respondent</td><td rowspan="3">10 552</td><td rowspan="3">8 855</td><td>Made up of a sole LU in U(LU)</td><td colspan="2">7 420</td></tr>
<tr><td rowspan="2">Made up of more than one LU in U(LU)</td><td>All LU are respondent</td><td>345</td></tr>
<tr><td>There is at least one LU with unknown R&D data</td><td>1 090</td></tr>
<tr><td>Non respondent</td><td>1 007</td><td colspan="4">894</td></tr>
<tr><td>Total</td><td>S(LU): 11 559</td><td colspan="4">S(EP): 9 749</td></tr>
</table>

Source: Ministry for Higher Education, Research and Innovation – 2015 R&D survey

6 The French national statistical institute.
7 The Frascati Manual is a manual from the OECD which gives the guidelines for collecting and reporting data on reasearch and experimental development.
8 Community innovation survey.
9 In the R&D survey, a large legal unit is a unit whose last known internal expenditure on R&D (BERD) is higher than 400 k€.

data are unknown either because it is not in the sample *S(LU)* (2 047 legal units), or because it did not answer to the survey (97 legal units). In the rest of the paper, an enterprise is considered respondent if at least one of its legal units has answered the R&D survey; otherwise, the enterprise is considered non respondent.

The second reason why data are not directly available for all the legal units belonging to an enterprise in *S(EP)* and to *U(LU)* is the existence of "grouped" responses in the survey, i.e. answers which are not about a sole legal unit only. Actually, in some situations, it was decided that some correspondents could answer for several legal units, independently of the concept of enterprise or group, to reduce its response burden.[10] These "grouped" responses, about a hundred or so each year, are matched with a "response outline" composed of several legal units. This kind of answer must be treated in a particular way to get the data for the legal units in the "response outline" and belonging to different enterprises of the sample *S(EP)*. In the 2015 survey, there are exactly 89 "grouped" responses (barely 1% of all the survey responses – cf. Table 2) which relate to 246 legal units and which represent 5.8 billion Euros of internal expenditures on R&D (BERD) (i.e. 18.3% of the total for the business sector). Among these 89 "grouped" responses, 29 mix several enterprises: they relate to 83 legal units and 63 enterprises.

Table 2 "Grouped" responses

Respondent population		Number	Legal unit level	Enterprise level
Non "grouped" responses		10 306	10 306	8 732
"Grouped" responses	*A sole enterprise*	60	163	60
	Several enterprises	29	83	63
Total		10 395	10 552	8 855

Note: Actually, this number is higher (8 782) because some enterprises have legal units in a "grouped" response and other ones not in "grouped" responses.
Source: Ministry for Higher Education, Research and Innovation – 2015 R&D survey

In conclusion, it is not so easy to compute R&D data for all the enterprises of the sample *S(EP)* from the data collected for the legal units sample *S(LU)*. Post-collection treatments are necessary.

1.3 The estimation of the internal expenditure on R&D (BERD) at the enterprise level

Within the framework of this study, we will consider only the internal expenditure on R&D (BERD) as R&D data. As seen in the previous paragraph, some enterprises in the sample *S(EP)* are made up of legal units in the target population *U(LU)* but whose BERD is unknown. We have to estimate this expenditure at the legal unit level to have a BERD estimator of good quality for such enterprises.

As said before, the "grouped" responses have to be treated in a particular way. As a consequence, firstly, we will concentrate on legal units whose BERD is unknown and which are not in a "grouped" response.

1.3.1 Estimation of the BERD for legal units except the ones in a "grouped" response

All the legal units in the target population *U(LU)* are likely to perform R&D but some of them do not. Unfortunately, this feature is not available for the legal units which are not in the sample S(*LU*) and for

[10] This may be due to strong collaboration on R&D activities between several legal units which may be close geographically, or to a unified accounting system (for two legal units in the same group for example).

the non-respondents. Then, a first step is to model the probability of performing R&D for each legal unit of the target population *U(LU)*. This modelling is based on all the respondent legal units except the ones in a "grouped" response (10 306 for the 2015 R&D survey – cf. Table 2), whether they have answered positively or not.[11] We distinguish four respondent sub-populations:

- the large legal units (QG[12] and exhaustive QS[13] – Large LU),
- the new legal units, i.e. the legal units which appear for the first time in the target population *U(LU)* (QS new – New LU),
- the legal units in the non-exhaustive stratum whose BERD in 2014 is unknown,
- the legal units in the non-exhaustive stratum whose BERD in 2014 is known.

For the large legal units, as the non-response and negatively response rates are low (respectively 0.15% and 3.94%), we assume that the probability of performing R&D, noted P(BERD>0) is equal to 1. For the other three respondent sub-populations, we estimate this probability thanks to a logistic regression. For each sub-population, the model is the following:

$$logit\left[P\left(BERD>0\right)\right] = \beta_0 + \sum_{k=1}^{K} \beta_k X_k + \epsilon, \tag{1}$$

where:

- β_0 is the intercept,
- $\beta_1, ..., \beta_K$ are the coefficients related to the K explanatory variables $X_1, ..., X_K$
- ϵ is the error term.

The explanatory variables for the different sub-populations are mentioned in Table 3. For the sub-population "legal units in the non-exhaustive stratum whose BERD in 2014 is known", there is an additional variable in the model: a dummy variable which equals 1 if the legal unit answered positively in the 2014 survey (BERD>0), 0 otherwise (BERD=0). Then, we estimate the coefficients for each of the three models by using the corresponding respondent legal units. These estimates allow us to estimate the probability of performing R&D for each legal unit whose BERD is unknown and which is not in a "grouped" response. From this estimated probability, we define a "performing R&D" dummy variable, noted IR&D, as follows:

$$I_{R\&D} = \begin{cases} 1 \, if \; P\left(\widehat{BERD>0}\right) > 0.5. \\ 0 \, otherwise \end{cases} \tag{2}$$

For some of the legal units (629), we cannot compute such a dummy variable because there are missing values for some explanatory variables. In this case, we assume that the "performing R&D" dummy variable equals 0.

Afterwards, a second step is to estimate the internal expenditure on R&D (BERD) for the legal units for which the "performing R&D" dummy variable equals 1 (i.e. $I_{R\&D}=1$). If the legal unit answered positively one of the previous R&D surveys between 2009 and 2014, we assume that the BERD in 2015 equals the last BERD known and corrected by price changes. Otherwise, the BERD is estimated by using a linear regression based on the legal units which answered positively the 2015 R&D survey (8 169[14] legal

[11] A legal unit answers positively the 2015 R&D survey if it has performed R&D in 2015 (BERD>0). Otherwise, it answers negatively (BERD=0).

[12] General questionnaire: questionnaire for legal units whose BERD exceeds 2 000 k€.

[13] Simplified questionnaire. Exhaustive QS are intended for legal units whose BERD exceeds 400 k€ but not 2 000 k€.

[14] Actually, the model is based on 7 609 legal units due to missing values and outliers.

Table 3 The explanatory variables[15] in the logistic regression models

<table>
<tr><th>Explanatory variable</th><th colspan="4">Values</th></tr>
<tr><td>Turnover (in k€)</td><td>[0;200[</td><td>[200;1 120[</td><td>[1 120;5 700[</td><td>[5 700;+ ∞[</td></tr>
<tr><td>Number of employees (headcount)</td><td>[0;2]</td><td>[3;9]</td><td>[10;32]</td><td>[33;+∞[</td></tr>
<tr><td>Share of exports in the turnover (in %)</td><td>0</td><td>]0;5[</td><td>[5;20[</td><td>[20;100]</td></tr>
<tr><td>Unit legal age</td><td>[0;2[</td><td>[2;12[</td><td>[12;23[</td><td>[23;+∞[</td></tr>
<tr><td rowspan="7">Business sector</td><td colspan="2">High</td><td colspan="2" rowspan="4">Technology industry</td></tr>
<tr><td colspan="2">Medium-high</td></tr>
<tr><td colspan="2">Medium-low</td></tr>
<tr><td colspan="2">Low</td></tr>
<tr><td colspan="4">Primary sector, energy, construction industry</td></tr>
<tr><td colspan="4">Services sector (excluding R&D)</td></tr>
<tr><td colspan="4">R&D (division 72 of the NACE Rev.2)</td></tr>
<tr><td rowspan="3">Part of business group</td><td colspan="4">Independent legal unit</td></tr>
<tr><td colspan="2" rowspan="2">Part of</td><td colspan="2">a French group</td></tr>
<tr><td colspan="2">a foreign group</td></tr>
<tr><td>Region of location</td><td colspan="2">Ile-de-France</td><td colspan="2">Other regions</td></tr>
<tr><td>Applying for R&D tax credit</td><td colspan="2">Yes</td><td colspan="2">No</td></tr>
</table>

Source: Ministry for Higher Education, Research and Innovation

units – cf. Table 4). In the corresponding model (model (2)), the variable to be explained is log(BERD) and the explanatory variables selected by using a stepwise procedure (the chosen significance threshold is 5%) are the following ones:

- the logarithm of the following continuous variables: turnover, number of employees and unit legal age;
- the share of export activities in the turnover;
- dummy variables corresponding to each value of the business sector, to the location in the region Ile-de-France and to the variable "Applying for R&D tax credit" (cf. Table 3);
- a "not large legal unit" dummy variable which equals 1 if the legal unit is not a large one, 0 otherwise.

[15] The continuous variables are turned into categorical ones with four values by using the quartiles.

Table 4 The number of legal units (LU) except the ones in a "grouped" response

Sub-populations		Respondent population		Non-respondent or non-sampled population			
					$I_{R\&D}=1$		
				$I_{R\&D}=0$	BERD estimated by		Estimated BERD (k€)
		DIRD=0	DIRD>0		previous BERD	linear regression	
Large LU		156	3 800	0	5	1	10 053
New LU		685	1 830	86	18	246	34 122
Non-exhaustive stratum	LU with unknown BERD in 2014	770	1 115	6 534	3 537	2 141	975 502
	LU with known BERD in 2014	526	1 424	733	2 069	43	464 011
Total		2 137	8 169	7 353	5 629	2 431	1 483 687

Source: Ministry for Higher Education, Research and Innovation – 2015 R&D survey

1.3.2 Estimation of the BERD for legal units in a "grouped" response

As seen in the paragraph 1.2., 29 "grouped" responses mix several enterprises. As a consequence, it is necessary to estimate the BERD for the 83 legal units involved in these "grouped" response to be able to estimate the BERD of the 63 enterprises related to these units in the end (cf. Table 2).

To estimate the BERD for legal units in a "grouped" response, firstly, we use the same model[16] (model (2)) and the same observations as the ones used in the previous paragraph (§1.3.1.) to get a first estimate. Then, we compute the share related to each legal unit in the estimated total BERD of the "grouped" response. Finally, we get the final estimate of the BERD for each legal unit by multiplying the corresponding share by the collected BERD of the "grouped" response which remains the same before and after estimation.

In conclusion, we have managed to get a value, collected or estimated, for the BERD of each legal unit of the target population *U(LU)*. As a consequence, we can compute a BERD for each enterprise of the sample *S(EP)*. To obtain an estimate of the total of the BERD from these values available at the enterprise level, we have now to determine a weight for each enterprise in the sample *S(EP)*.

2 HOW TO GET AN ESTIMATE OF THE TOTAL BERD FROM THE RESPONDENT ENTERPRISES?

2.1 The respondent enterprises

In this paragraph, we will naturally estimate the total BERD by considering the 8 855 respondent enterprises (cf. Table 1), i.e. the ones whose at least one unit legal answered the 2015 R&D survey. The whole of these enterprises is noted $S(EP)_r$. Thanks to the paragraph 1, a BERD value can be computed for each of them by adding the BERD, collected or estimated (cf. Table 5), of each of their legal units.

[16] Actually, the explanatory variables are not exactly the same: the "not large legal unit" dummy variable is substituted by two dummies, namely a "QG" dummy and an "exhaustive QS" one, because only 2 "grouped" responses are not in this sub-population.

Table 5 The legal units of the respondent enterprises

	Number of respondent enterprises	Number of legal units in the respondent enterprises (BERD (k€))		
No estimated BERD for all the legal units in the enterprise	7 688	8 124 (15 303 200)		
Estimated BERD for at least one of the legal unit in the enterprise	1 167	Collected BERD		2 345 (11 176 213)
		Estimated BERD	"Grouped" responses connected to several enterprises	83 (3 459 533)
			Non-sampled or non-respondent	2 143 (367 006)
TOTAL	$S(EP)_r$: 8 855	12 695 (30 305 952)		

Note: The total is not weighted for the BERD.
Source: Ministry for Higher Education, Research and Innovation – 2015 R&D survey

To obtain the total BERD, we just have to determine a weight for each respondent enterprise. We compute this set of weights by using generalized weight share method (GWSM) (Lavallée, 2007).

2.2 Generalized weight share method (GWSM)

In the framework of the generalized weight share method (GWSM), we distinguish two populations: the collection population, here *U(LU)*, and the population of interest, here *U(EP)*. The GWSM allows to obtain an unbiased estimator of the variable of interest, here the total BERD, from the data available at the interest population level and the weights of the units sampled in the collection population, i.e. here the legal units in the sample *S(LU)*. In this study, only the weights of respondent legal units, i.e. the legal units in $S(LU)_r$, are considered because total non-response is handled here through reweighting.

The estimator of the total BERD related to the GWSM, noted $\widehat{BERD\_tot}_{GWSM}$ is expressed as follows:

$$\widehat{BERD\_tot}_{GWSM} = \sum_{EP_i \in S(EP)_r} pond\_EP_i^{GWSM} \times BERD\_EP_i, \tag{3}$$

where:

– $BERD\_EP_i$ is the BERD computed for the enterprise EP_i by adding the BERD of each of its legal units which are in the collection population *U(LU)*:

$$BERD\_EP_i = \sum_{LU_{k,i} \in EP_i \cap U(LU)} BERD\_LU_{k,i}, \tag{4}$$

– $pond\_EP_i^{GWSM}$ is the weight for the enterprise EP_i computed thanks to the GWSM. It is expressed as follows:

$$pond\_EP_i^{GWSM} = \sum_{LU_{k,i} \in EP_i \cap S(LU)_r} \theta_{k,i} \times pond\_LU_{k,i}. \tag{5}$$

Pond_LU$_{k,i}$ is the final weight of the legal unit k of the enterprise *i*, i.e. the weight got after reweighting to handle total non-response.

The values taken by the coefficients $\theta_{k,i}$ depend on the version of the used GWSM. In this paper, we consider the two following versions:

- the GWSM with classical links: in this version, the weight of an enterprise is based on the number of its legal units which are in the collection population *U(LU)*:

$$\forall LU_{k,i} \in EP_i \cap S\left(LU\right)_r, \quad \theta_{k,i}^{classical} = \theta_i^{classical} = \frac{1}{\sum_{LU \in EP_i \cap U(LU)} 1}, \tag{6}$$

- the GWSM with links weighted by the BERD: in this version, the BERD of legal units is introduced as a weight in the calculation of the coefficients $\theta_{k,i}$:

$$\forall LU_{k,i} \in EP_i \cap S\left(LU\right)_r, \quad \theta_{k,i}^{weighted\_BERD} = \frac{BERD\_LU_{k,i}}{BERD\_EP_i} \tag{7}$$

Then, we have built two new estimators of the total BERD, $\widehat{BERD\_tot}_{GWSM}^{classical}$ and $\widehat{BERD\_tot}_{GWSM}^{weighted\_BERD}$. In the next paragraph, we will compare them with the current estimate of the total BERD, noted $\widehat{BERD\_tot}_{LU}$, got from the BERD collected at the legal unit level and which can be written as:

$$\widehat{BERD\_tot}_{LU} = \sum_{k \in S(LU)_r} pond\_LU_k BERD\_LU_k, \tag{8}$$

and we will try to identify the best of them.

3 WHICH ESTIMATOR AT THE ENTERPRISE LEVEL TO RETAIN?

3.1 Comparison of the two GWSM estimators

To assess the quality of our estimators, we would like to compare them to the current estimator $\widehat{BERD\_tot}_{LU}$. But the computation of the GWSM with BERD-weighted links estimator of the BERD mechanically gives the same result as the current one. Another idea is to compare our two GWSM estimators to the real totals, i.e. the totals on the population *U(EP)*, for the BERD[17] and two variables available in the French business register SIRUS (turnover and headcount). In this paper, we will assess the quality of our two GWSM estimators by computing the relative deviation from the real total for these three variables (cf. Table 6).

Both number of enterprises and legal units are underestimated by the two GWSM estimators, but the estimations of the totals for the two non-R&D variables (turnover and headcount) and for the BERD are quite close to the real ones, especially in the case of the GWSM with BERD-weighted links estimator. As a consequence, the latter seems to be the best one with the "relative deviation" criterion. However, further studies would be necessary to consolidate this first result. For example, we could carry out several simulations to compute several values for both GWSM estimators and, then, to deduce the bias for each estimator. In this paper, we focus only on the "relative deviation" criterion and, so, we decide to choose the GWSM BERD-weighted links estimator as the best one to compute R&D statistics at the enterprise level.

17 Actually, about the BERD, we can only compute a pseudo-real total on the population *U(EP)* thanks to estimations conducted in the first paragraph.

Table 6 GWSM estimations based on the enterprise population and corresponding relative deviations from the real totals

	U(EP)	*S(EP)$_r$*, GWSM with classical links		*S(EP)$_r$*, GWSM with BERD-weighted links	
	Real total	Estimated total	Relative deviation from the real total	Estimated total	Relative deviation from the real total
Number of enterprises	21 466	19 106	–11%	19 254	–10%
Number of legal units	25 965	23 255	–10%	23 838	–8%
Turnover (B€)	2 038	1 976	–3%	2 066	1%
Headcount (thousands)	5 740	5 416	–6%	5 509	–4%
BERD (M€)	31 423	30 640	–2%	31 756	1%

Source: Ministry for Higher Education, Research and Innovation – 2015 R&D survey; INSEE – SIRUS

Another way to judge the quality of this new estimator is to understand how relevant it is in terms of economic analysis in comparison to an analysis at the legal unit level.

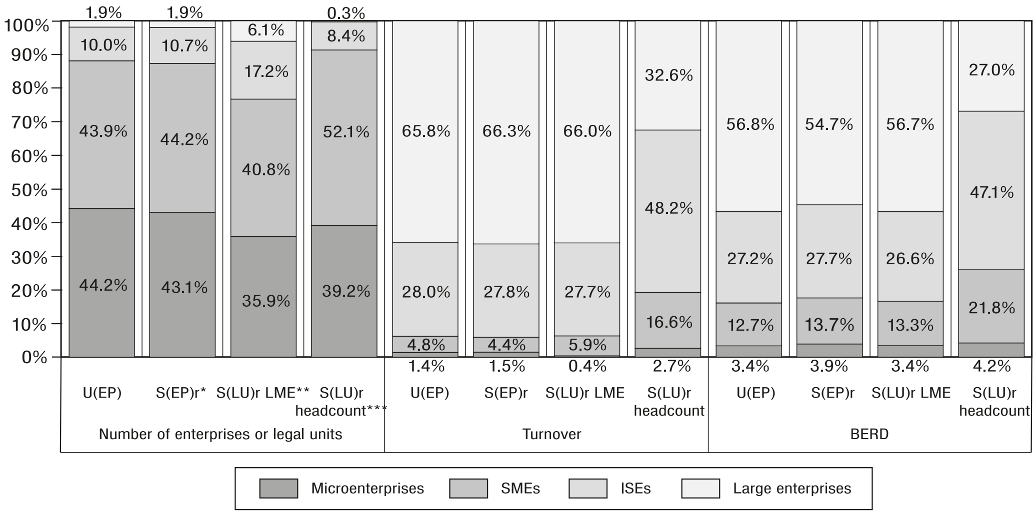
3.2 Comparison with the legal unit level: which level is better for an analysis by business category?

As stated in the previous paragraph, the GWSM BERD-weighted links estimator gives the same total BERD as our current estimator computed at the legal unit level. However, another interesting and topical issue is the analysis by business category. Indeed, the law (LME)[18] specifies now four categories of business for the purposes of statistical and economic analysis, using headcount, turnover and total balance sheet of the enterprise: microenterprises, small and medium-sized enterprises (SME), intermediate-sized enterprises (ISE) and large enterprises. Before the new French definition of enterprise (LME), legal units were only categorized by their size in terms of headcount, with the same threshold than in the new definition. Moreover, data have already been published on LME business categories from the legal unit level: for example, the BERD of SMEs is the sum of the BERD of all legal units belonging to a SME. With this method, we obtain the number of legal units in each enterprise category but not the number of enterprises. So, it's interesting to compare the results got from our new estimator at the enterprise level (the GWSM BERD-weighted links estimator) to those got at the legal unit level, in terms of breakdown by business category (cf. Figure 1).

With the new French definition of enterprise, we completely change what we say about the breakdown between business categories: with the headcount definition on legal units we underestimated the share of large enterprises in the total BERD, and with the LME definition on legal units the shares in the total BERD were well estimated but we could not say anything in terms of number of enterprises. Thanks to our new methodology, we can now compute estimators directly based on the enterprises in our population, for example the number of enterprises carrying out R&D activities.

[18] Article 51 of the French law on the modernisation of the economy (LME), 2008.

Figure 1 Breakdown of some variables by business category at the enterprise level and at the legal unit level



Notes: * Estimators computed with the GWSM with BERD-weighted links on the respondent enterprises *S(EP).*
** Estimators at the legal unit level, LME definition of business categories.
*** Estimators at the legal unit level, business categories defined only by headcount thresholds.
Reading note: There are 44.2% microenterprises in *U(EP)*, whereas we estimate there are 35.9% legal units belonging to a microenterprise ($S(LU)_r$ LME).
Source: Ministry for Higher Education, Research and Innovation – 2015 R&D survey; INSEE – SIRUS

CONCLUSION

The data collection at the legal unit level is still interesting and convenient, and the "grouped" responses and non-respondent legal units belonging to a respondent enterprise prevent us to easily rebuild the enterprise data. Different models were therefore used to rebuild a total BERD for each enterprise. This allows us to compute more relevant estimators at this enterprise level, especially in terms of business categories, thanks to the generalized weight share method (GWSM) which generates new sets of weights.

To consolidate our results, it would be necessary to assess more thoroughly the quality of our GWSM BERD-weighted links estimator, for example by carrying out simulations based on different samples to compute different expected values of our total BERD estimator and its bias. Then, we could use our new set of weights at the enterprise level to estimate the totals of other variables, such as R&D personal or researchers, and check the quality of those estimators.

References

OECD. *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities.* Paris: OECD Publishing, 2015.

LAVALLÉE, P. *Indirect sampling.* Springer, 2007.

HAAG, O. Profiling: a new and better way to apprehend the globalization. *Q2016 Conference*, Madrid, 2016.

HAAG, O. The French business registers system: How to improve the quality of the statistics by combining different statistical units. *The Fifth International Conference on Establishment Surveys* (ICES-V), Geneva, 2016.

DUNNING, J. H. AND LUNDAN, S. M. The Internationalization of Corporate R&D: a Review of the Evidence and Some Policy Implications for Home Countries. *Review of Policy Research*, 2009, No. 26, pp. 13–33.

FLORIDA, F. The globalization of R&D: Results of a survey of foreign-affiliated R&D laboratories in the USA. *Research Policy*, 1997, 26, pp. 85–103.

MONCADA-PATERNO-CASTELLO, P., VIVARELLI, M., VOIGT, P. Drivers and Impacts in the Globalization of Corporate R&D: the European Experience. *Institute for Prospective Technological Studies Working Paper Series on Corporate R&D and Innovation*, Luxembourg: Publications Office of the European Union, 2011, No. 3.

The Improvement of Response Rates and Data Quality of Direct Business Surveys by Centralized Data Collection Approach: the ISTAT Experience

Giampaola Bellini [1] | *ISTAT – Italian National Statistical Institute, Rome, Italy*
Francesca Monetti [2] | *ISTAT – Italian National Statistical Institute, Rome, Italy*
Pasquale Papa [3] | *ISTAT – Italian National Statistical Institute, Rome, Italy*

Abstract

In April 2016 ISTAT (Italian National Statistical Institute) started a corporate restructuring process that interested all the statistical production structures and that led to a completely renewed organizational setup. Before the above mentioned reorganization, the statistical processes were organized according to the classical 'stovepipe' model, that involved independent, non-integrated, statistical processes including all the necessary skills: statisticians, information technology experts, thematic experts, methodologists.The new model restricts the production processes only to the thematic experts, while all the "cross" expertise isare all assigned to specialized structures. The main advantage of the new setup concerns the overall system efficiency, while the main disadvantage concerns the increased fragmentation of the production processes.

Before the restructuring process, response rates in economic structural surveys were quite low and unsatisfactory. After two years from the introduction of the new organization the medium response rate increased from 48.8 to 59.5 per cent for structural surveys and from 59.0 to 79.0 for short-term surveys. At the same time, the duration of the data collection periods for structural surveys reduced from 152 to 115 days.[4]

Keywords

Data collection, response rates, process efficiency, official statistics, business stastistics, quality of statistics, non-response, statistical survey

JEL code

C81

[1] ISTAT, Via Cesare Balbo, 16 – 00184 Rome, Italy. Paragraphs: 1.1, 2.4, 2.5, 2.6, 2.8.2.
[2] ISTAT, Via Cesare Balbo, 16 – 00184 Rome, Italy. Paragraphs: 1.2, 2.1, 2.2, 2.3, 2.7.
[3] ISTAT, Via Cesare Balbo, 16 – 00184 Rome, Italy. Corresponding author: email: papa@istat.it, phone (+39)0646734055. Paragraphs: 1 (Introduction), 2.8.1, 3 (Conclusion).
[4] This contribution was presented at *European Conference on Quality in Official Statistics* (Q2018), Krakow, 26–29 June 2018.

INTRODUCTION

The Response rates (*rr*[5]) in economic direct surveys carried out in Italy were traditionally low. Low *rr* are partially explained by structural characteristics of Italian economic productive system, that includes a very high number of medium and small-size companies. In 2015, of the total of 4 338 085 active companies only 3 666 had more than 250 employees, just 0.08 per cent. The total number of employees was 16 289 875 of which 3 583 121 belonging to enterprises having more than 250, representing a quota of 22%. Traditionally, in Italy, medium and small size companies showed lower *rr* than the bigger ones. An explanation of different *rr* corresponding to different dimensional classes depends on the characteristics of the Italian statistical law that imposes an obligation to provide data for all the companies involved in statistical surveys but it applies penalties just for a reduced set of companies, identified according to dimensional variables (number of employees and/or turnover). During the last two years Istat experienced a clearly increasing trend in *rr* both in structural and short-term economic surveys. The increase of the *rr* was normally associated with a significant reduction of the data collection period. Particularly for main structural economic surveys[6], the results show (Table 1) that generally speaking the *rr* increased by about 11 percentage points (*pp*), whereas for short-term surveys by about 20 *pp*. The comparisons refer to surveys on enterprises carried out before and after the Centralized Data Collection (CDC) management was implemented in Istat. Notably for short-term surveys the comparison was carried out considering 1st quarter 2016 and 1st quarter 2018. In the above mentioned framework the main aim of the article is to point out the effects of CDC on response rates of both structural and short-term economic surveys.

Table 1 General response rates (rr) for main structural and short-term surveys pre and post CDC implementation

Type of business survey	Management type	Total survey units (*)	General *rr* (average for short term)
Structural (**)	Pre CDC	264 698	48.8
	Post CDC ***	231 681	59.5
Short-term (**)	Pre CDC	55 512	59.0
	Post CDC	57 667	79.0

Notes: * Each unit can be included in one or more surveys. ** The main structural surveys and a selection of short-term surveys are included. See Tables 2 and 4 for the complete lists. *** Considering last concluded survey.
Source: Elaboration on data extracted from Business Statistical Portal

1 TRENDS OF RESPONSE RATES IN ECONOMIC SURVEYS AFTER STARTING THE MODERNIZATION PROGRAM

1.1 Response Rates in Structural surveys

Among the structural surveys, some responded very positively to the new organizational scheme, that are not only the ones characterized by a low *rr* as the Community innovation survey (CIS) that increased by 15 *pp* in 2016 starting from 53 percent in 2014, but also surveys with an already satisfactory *rr* such as Survey on enterprise accounting system (SBS – Structural Business Statistics Regulation), that increased 8 *pp*, starting from 68 in 2014 (Table 2).

[5] In this document *rr* are calculated at the end of the data collection phase but before activating the procedures for integrating the missed responses, which may vary from survey to survey, and before making any integration with data from administrative sources.

[6] Community innovation survey (CIS), Statistics by product (Prodcom), Small and medium size enterprise survey – SME (including professional and artistic activities) (SBS), Survey on information and communication technology in enterprises (ICT), Survey on enterprise accounting system (SBS), Survey on Research and Development in enterprises (R&D), Updating of the statistical register of economic units ASIA – Local units, Survey on the activities of foreign controlled enterprises resident in Italy (Inward Fats), Survey on foreign affiliates activities abroad controlled by Italy (Outward Fats).

Table 2 Structural surveys, response rates and data collection periods lengths

Surveys	Reference year	Total units	Response rate (%)	Response rate difference (%) (*)	Data collection length (total days – d)	Data collection length difference (d) (*)
CIS	2016	32 018	68.1	15.1	143	–92
Prodcom	2017	39 799	56.2	11.6	124	2
SME – SBS	2016	74 207	43.5	11.0	99	–118
ICT	2017	32 255	67.0	5.2	66	–27
SBS	2016	10 558	76.4	8.4	139	–4
R&D	2015	17 977	76.5	–1.0	89	–48
Updating the statistical register of economic units ASIA – Local units	2016	10 536	80.4	–3.8	62	–51
Inward Fats	2015	7 791	74.4	24.2	120	–21
Outward Fats	2016	6 326	69.8	10.2	193	24

Note: * Comparison was carried out between last concluded survey and the last run before CDC introduction.
Source: Elaboration on data extracted from Business Statistical Portal

Regarding the length of the survey, we can highlight the Small and medium size enterprise accounting system survey – SME (including professional and artistic activities – SBS Regulation) case that was run in 118 days (*d*) less than the previous editions, and the Community innovation survey (CIS) with 92 *d* less; several other surveys were conducted in shorter time, with more than 20 *d* less than usual (Table 2). The analysis of *rr* related to enterprise dimension, enhances the effect of several factors. Table 3 below shows that a relevant increase was registered for SBS (8.9 *pp*) and ICT (9.4 *pp*) surveys, where the enterprises involved with at least 250 employees are more than 3 000. In both cases the increase was higher than the general *rr* variation of each survey (8.4 *pp* for SBS and 5.2 pp for ICT), meaning that the impact of the new organizational scheme garantees effectiveness of the activities run particularly on the larger units. Moreover, the response rates for enterprises having at least 250 employees of these two surveys before the reorganisation was lower than the ones registered for Prodcom and Outward Fats. On the other

Table 3 Structural surveys, response rates for enterprises having at least 250 employees

Surveys	Response rate (*)	
	Total (%)	*pp* difference (*)
SBS	86.2	8.9
Prodcom	89.0	–0.5
ICT	91.6	9.4
Outward Fats	90.3	4.9

Note: * Comparison was performed between last concluded survey and the one run before CDC introduction.
Source: Elaboration on data extracted from Business Statistical Portal

hand, for Prodcom and Outward Fats the variation of the *rr* for enterprises with at least 250 employees was quite different from the previous ones with a decrease of 0.5 *pp* and an increase of 4.9 *pp* respectively and definitely lower than the variations recorded for the survey as a whole (11.6 and 10.2 *pp* respectively). In those two cases, the number of units involved with this dimension is not very high (around 1 200 and 500, respectively) and the *rr* for this specific group was already high before reorganisation (88.9 and 85.4 percent, respectively).

1.2 Response Rates in Short-term surveys

Tables 4, 5 and 6 show the results obtained for the first quarter 2018 in terms of *rr* after CDC introduction for a selected set of business short-term surveys. Since these surveys are characterized by a continuous DC (Data Collection) process, the comparison was made at the end of the useful period, introduced with the new sanctioning procedure (paragraph 2.7). Compared with the first quarter 2016, the *rr* shows a positive average variation of 20.0, particularly relevant is the increase of 28 *pp* registered for the Monthly survey on retail sales (MRS). The Monthly survey on industrial production (IPI) and the Monthly survey on producer prices for industrial products sold in the domestic market (PPID) also show significant increases of 20.1 *pp* and 11.9 *pp*, respectively (Table 4).

Table 4 Short-term surveys, average response rates – I quarter, years 2016 (pre CDC) and 2018

Survey	I quarter 2016 (%)	I quarter 2018 (%)	*pp* difference
MRS	38.3	66.3	28.0
IPI	58.8	78.9	20.1
PPID	79.9	91.8	11.9

Source: Elaboration on data extracted from Business Statistical Portal

It is also noted that the *rr* of the short-term surveys in 2018 compared to 2016, despite the more stringent tolerance times envisaged by the current sanctions system, are significantly higher at the date of May 2018, particularly for the surveys MRS and PPID.

Table 5 Short-term surveys, average response rates for enterprises having at least 100 employees – I quarter, years 2016 (pre CDC) and 2018

Survey	I quarter 2016 (%)	I quarter 2018 (%)	*pp* difference
MRS	51.9	78.1	26.2
IPI	60.2	84.6	24.4
PPID	76.6	94.2	17.6

Source: Elaboration on data extracted from Business Statistical Portal

The increase concerning enterprises having at least 100 employees (Table 5), highlighted by the average *rr* variation from 62.9 to 85.6 percent, associated with an average reduction in the number of enterprises virtually subject to penalties of 63.4 percent according to the new management criteria, is considerably positive (Table 6).[7]

[7] Data used to calculate the response rates in Tables 4 and 5 may not coincide with those used for calculating the indicators transmitted to Eurostat and published at national level; due to the fact that even if there is a monthly deadline for sending the data, the enterprises can still provide the information throughout the month also referred to previous periods, in this case the data are used for the review of indicators. For example, the response rates calculated for the IPI survey near the Press Release at national level for January 2016 and January 2018 are equal to 67% and 81%, respectively, with a positive variation of 14%.

Table 6 Business short-term surveys, number of enterprises virtually subject to penalties – I quarter, years 2016 and 2018*

Survey	Enterprises subject to penalties		
	I quarter 2016	I quarter 2018	% decrease
MRS	136	55	–56.6
IPI	1 334	532	–60.1
PPID	378	111	–70.6

Note: * Number of enterprises virtually subject to penalties is calculated at the end of the 'useful data deadline'.
Source: Elaboration on data extracted from Business Statistical Portal

Given the results mentioned, the main objective of the present document is to point out the causes that explain both the increasing trend in response rates and data collection period reduction. As pointed out in the following text, the main causes concern organizational set-up solutions of the data collection processes, that involved increasing efficiency and standardization. All the results were obtained thanks to the synergies established among DC structures, production structures and IT structure.

2 INNOVATIONS INTRODUCED IN THE FIELD OF DC IMPLEMENTATION

2.1 Organizational set-up of DC implementation

During 2016 the Italian National Statistical Institute (ISTAT) launched a wide modernization programme whose main objective was to enrich the supply and quality of the information produced, improving the effectiveness and efficiency of the statistical processes. Istat designed and implemented a new organizational set-up that was characterized by the centralisation of all the support services, that were clearly separated from statistical production. The most important innovation subsisted in the creation of the new Central Directorate for Data Collection that was characterised by a very high degree of specialization of activities and Human Resources. In fact it included four Divisions specialised in the following areas: 1) Division for design of data collection tools; 2) Division for data collection organization; 3) Division for implementation of data collection from direct surveys; 4) Division for integration of administrative sources and registers.

In the reorganization of data collection the goal of specialization was pursued concentrating a series of activities typical for survey's implementation in a single Division. A further internal subdivision concerned the type of responding units involved (businesses, households and individuals, farms, public and private institutions, others). The integration of data collection implementation processes, previously run independently, promoted standardization, with a view to optimizing and increasing efficiency. As a consequence, several process innovations were implemented.

2.2 Harmonized survey lists management

The preparation of the survey lists was standardized and generalized, by means of a new procedure involving two successive steps of treatment: i) verification of the elegibility of the units included in the survey samples, in order to define the correct and updated lists of units to be involved in the survey. These units receive the informative letter, signed by Istat's President, announcing the start of the survey. Eligibility is assessed taking into account possible recent business transformation events, insolvency proceedings, registrer modifications and economic activity variations; ii) normalization of the mailing list, verifying for each unit the completeness of register information useful for the correct delivery and integrating possible gaps.

2.3 Standardization of the contact modalities

The following standards were adopted: i) single centralized access point both for the data capturing systems (by means of Business Statistical Portal) and for the incoming contacts (free of charge inbound Contact Center) by telephone or by email; ii) system of harmonized standard answers to be used in order to provide efficient and timely resolutions to requests coming from units on non-thematic and recurring thematic issues. The requests are addressed to centralized inbound Contact Center service or directly to Istat DC offices.

2.4 Strict schedulation for formal end informal communications

The data collection implementation requires the definition of a strict timetable for the management of the formal and informal communications addressed to the units involved in the surveys. The following timetable has been adopted, following different approaches for structural and short-term surveys.

Table 7 Timetable of formal and informal communications adopted by structural and short-term surveys

<table>
<tr><th rowspan="2">TYPE OF SURVEY</th><th rowspan="2" colspan="2">DC START</th><th colspan="8">FIELD DATA COLLECTION</th></tr>
<tr><th colspan="4">Survey reminders pre-deadline</th><th>Deadline for data submission</th><th colspan="3">Survey reminders post-deadline</th></tr>
<tr><td>STRUCTURAL</td><td rowspan="2">Sending informative letter</td><td rowspan="2">Sending ordinary email to survey contact persons</td><td>First reminder by certified email (halfway survey period)</td><td>Second reminder by certified and ordinary email (around a month before deadline)</td><td>Extra reminder by certified and/or ordinary email (for surveys with low rr or short data collection period)</td><td>Telephone recall (about twenty-one to seven days before deadline)</td><td>Informative letter deadline (data capturing system dosure)</td><td>–</td><td>–</td><td>–</td></tr>
<tr><td>SHORT-TERM</td><td colspan="4">Reminder by ordinary email (about two days before informative letter deadline)</td><td>Date of informative letter deadline</td><td>Reminder by certified email (about two to ten days after deadline)</td><td>Telephone recall (about five to ten days after deadline)</td><td>Reminder by ordinary email to survey contact persons (only to enterprises subject to penalities)</td></tr>
</table>

Source: ISTAT

The massive submissions are carried out through a specific software application (named Archiflow) that allows the creation and sending by certified email, scheduling the starting time, of dynamic text messages; massive ordinary email dynamic text messages are managed by means of a proprietary Web application named MMM (Mail Massive Manager).

2.5 Procedures and tools for monitoring the data collection process

Automatic and generalized procedures were implemented in order to monitor the entire data collection process. The aim is implementing timely corrective actions to control non-respondents, such as extra reminders in addition to those already scheduled (by ordinary or certified email, or phone). These procedures allow cyclical (comparing to the previous period of the same year) and structural comparisons (same period of the previous year) on the basis of specific indicators (e.g. response rates). In this regard, particular relevance assumes the management tools provided by the back office of the Business Statistical Portal that allows a detailed analysis of the *rr* for territorial level, economic activity and for specific employees classes.

2.6 Data capturing and security systems

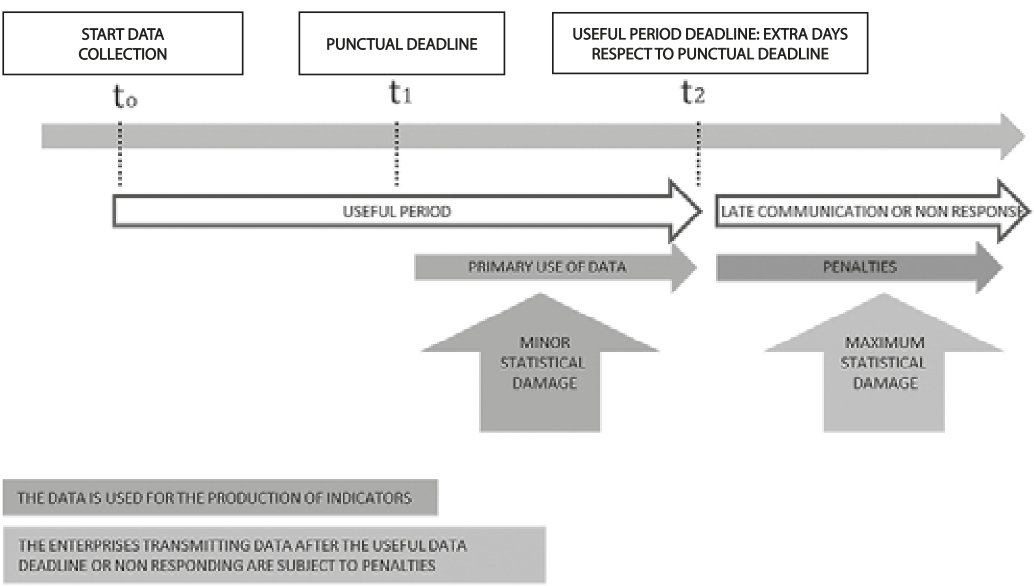
Data capturing takes place in a safe mode through generalized Web systems which allow the storage of raw data in a separate and centralized logical environment that allows monitoring of all deliveries to different recipients.

2.7 Harmonised penalties management procedure

The integrated approach to the CDC management allowed the generalization of the procedures used for the generation of the lists of the units subject to penalties. The lists are produced at the end of the DC period, after appropriate check of the most recent register information.

In particular, in the context of the short-term surveys an important innovation was introduced, aimed at redefining the procedure for the identification of units subject to penalties. The new procedure was implemented with the aim to produce timely and quality statistical information while trying to minimise the statistical damage charged on Istat; the statistical damage has been assessed on the basis of the response behaviour by the units involved in the surveys in relation to the phases and timings, as reported in the following Figure 1.

Figure 1 Penalties management criteria in short-term surveys



Source: ISTAT

2.8 Innovative tools and services supporting DC activities

2.8.1 The Business Statistical Portal

The introduction of the Business Statistical Portal in economic survey started in 2013 involving a reduced set of large companies (500 employees and more). In May 2018 about 350 000 companies for about 60 surveys are currently enabled to use the Portal. The implementation of the new system in the context of the economic surveys involved a new approach in the management of economic surveys that turned from "survey-centered" to "enterprise-centered". Main objectives of the Portal can be defined as follows: a) Streamline the operations required by respondents to fulfill their response obligations, with an overall reduction of the burden; b) Increase both ordinary and extraordinary (e.g. news) communications on the survey events and activities; c) standardize and harmonize data collection in order to increase efficiency at the system level.

2.8.2 Centralised inbound and outbound Contact center services

The new organization of Division for data collection implementation from direct surveys (DCI) also implies more specialization of managing the contacts with respondents. In particular, the outsourcing of the activity is entrusted to a specialized company in Contact Center (CC) services. The aim is pursuing progressive centralization of the support and assistance services addressed to the units involved in the surveys (inbound) and of telephone alert and reminders addressed to non-respondent units (outbound). The unique and coordinated management of the service guarantees strong standardization not only within each specific thematic sector but also among sectors, due to the increased transfer of the best practices from one sector to the other.

CONCLUSIONS AND FUTURE DEVELOPMENTS

The introduction of the new organizational model launched by ISTAT in 2016, which provides a specialized approach to the management of cross-cutting services and the creation of a new Department exclusively dedicated to the Data Collection has produced important results in terms of increasing response rates of economic business surveys, both structural and short-term. The increases are also associated to significant reductions in the data collection periods, especially in structural economic surveys. The results are independent of the platforms used for web data capturing and are extended to all types of surveys. Among the factors that most explain these increases should be considered the standardization of data collection processes that led to significant increases in efficiency. These efficiency gains can free up resources to be used in process and product innovation activities, in the quality of the outputs and to respond to new needs for statistical information expressed by users. Even in the presence of the above mentioned undoubted results, the new organization of the processes has also shown some critical issues that can be resolved in the medium term: i) resistance to change and increase in the conflict between transversal and production structures, mainly deriving from the "subtraction" of some activities that were traditionally managed within the production processes; ii) strong fragmentation of DC processes; iii) permanence of overlaps and doubts about "who does what" in the transversal structures and in particular Data collection.

The main challenges for the future concern the methods and the solutions to be adopted to consolidate the transition process towards the new model: i) development of integrated and generalized platforms for data capturing from units belonging to different sectors; ii) design and implementation of a unique generalized system of integrated management of surveys; iii) greater integration between inbound and outbound contact center services; iv) development of acquisition Portals to increase the efficiency of data collection processes from survey units belonging to different sectors; v) identify solutions to be applied at an organizational level in order to reduce the processes fragmentation, while respecting the principle of specialization and standardization of the activities involved.

References

EUROSTAT. *ESS Guidelines for the implementation of the ESS Quality and Performance Indicators* (QPI) [online]. 2014. <https://ec.europa.eu/eurostat/documents/64157/4373903/02-ESS-Quality-and-performance-Indicators-2014.pdf/5c996003-b770-4a7c-9c2f-bf733e6b1f31>.

EUROSTAT. *European Statistics Code of Practice* [online]. 2011. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-32-11-955>.

FAZIO, N. R, MURGIA, M., NUNNARI, A., *The business statistical portal: a new way of organizing and managing data collection processes for business surveys in Istat.* UNECE Conference of european statisticians, Seminar on Statistical Data Collection, Geneva, Switzerland, 25–27 September 2013.

ISTAT. *Mapping delle attività della DCRD nell'ambito dello schema concettuale di riferimento internazionale GSBPM*. Delibera D16 49 DIRM2017, 2017.

RIVAIS, L., ST-DENIS, M., LENSEN, S. *Centralising data collection at Statistics Canada*. UNECE Conference of european statisticians, Seminar on Statistical Data Collection, Geneva, Switzerland, 25–27 September 2013.

SARAIVA DOS SANTOS, P. AND MOREIRA, A. *Creating a data collection department: statistics Portugal's experience*. Seminar on Statistical Data Collection, Geneva, Switzerland, 25–27 September 2013.

SIGNORE, M. *GSBPM and other international standards MedStat training on GSBPM* [online]. Rome: ISTAT, 2017. <https://statswiki.unece.org/display/GSBPM/GSBPM+Training+Materials>.

SIGNORE, M. *GSBPM how to use and implement MedStat training on GSBPM* [online]. Rome: ISTAT, 2017. <https://statswiki.unece.org/display/GSBPM/GSBPM+Training+Materials>.

MARSKE, R. AND STEMPOWSKI, D. M. *Company-Centric Communication Approaches for Business Survey Response Management* [online]. Data Collection: Challenges, Achievements and New Directions, Statistics Canada Symposium, 2008. <https://www150.statcan.gc.ca/n1/en/pub/11-522-x/2008000/article/10983-eng.pdf?st=0-IwAKvz>.

UNECE STATISTICS WIKI. *Generic Statistical Business Process Model – GSBPM* [online]. Geneva, 2019. <http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model>.

International Conference *Mathematical Methods in Economics* (MME 2019)

Petra Zýková [1] | *University of Economics, Prague, Czech Republic*
Josef Jablonský [2] | *University of Economics, Prague, Czech Republic*

Mathematical Methods in Economics (MME) conferences have a very long history and tradition. They are one of the most important scientific events organized in the Czech Republic in the field of operational research, econometrics, mathematical economics, and related research areas. In 2019, the 37th *International Conference on Mathematical Methods in Economics* was organized in the city of České Budějovice on 11–13 September.[3] Except for the local organizer, which was the Faculty of Economics, University of South Bohemia in České Budějovice, leading organizers of MME conferences are the *Czech Society for Operations Research* (CSOR) and the *Czech Econometric Society*.

The total number of participants of this year's MME conference was more than 127 from the Czech Republic, China, Austria, Poland, and Slovakia. The programme started with an opening ceremony, where the head of the Organising Committee, Michal Houda, introduced the main programme and all facilities. After that, the first plenary session started with the interesting lecture *From AI and Data Science back to Operations Research and Financial Modelling* presented by the Vice President of the Austrian Operations Research Society, professor Ronald Hochreiter (Vienna University of Economics and Business). After the plenary session, the programme of the conference was divided into five parallel sessions. The total number of presentations was more than 80. All accepted papers are published in the Proceedings of the MME 2019. They have been submitted, as in the previous years, for indexing in the Web of Science database.

It has been a long tradition that a competition of PhD students for the best paper takes place during MME conferences. The competition is organized and honoured by the CSOR. All submitted papers were peer-reviewed and the papers with positive referee reports were further evaluated by the Programme Committee. Six best selected papers have been presented at the conference in a special session and the evaluation committee decided about the winners. Six papers, which were the best ones, were awarded after a conference dinner on the Kleť mountain. Petra Tomanová (University of Economics, Prague, Czech Republic) with her paper *Price Clustering Phenomenon* was the winner of the competition. Tomáš Rusý (Charles University, Prague, Czech Republic) with his paper *Interest Rate Modelling: Maximum Likelihood Estimation of One-Factor Short-Rate Models* ranked second. Petra Zýková (University of Economics, Prague, Czech Republic) ranked third with her paper *Dynamic Efficiency Analysis of German NUTS 2.* The remaining three award-winning contributions have been delivered by Anlan Wang (Technical University

[1] Faculty of Informatics and Statistics, Department of Econometrics, University of Economics, Prague, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: petra.zykova@vse.cz.
[2] Faculty of Informatics and Statistics, Department of Econometrics, University of Economics, Prague, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: jablon@vse.cz.
[3] More at: <*https://mme2019.ef.jcu.cz*>.

of Ostrava, Czech Republic), Michal Škoda (Czech University of Life Sciences Prague), and Xiaoshan Feng (Technical University of Ostrava, Czech Republic). Another plenary session closed the conference by Gustav Feichtinger (Vienna University of Technology) and his lecture *The Mathematics of Ageing*.

Organization of the conference was at a very high level. All sessions took place in the Faculty of Economics. The welcome evening took place on the ground floor of the university canteen with an accompanying music (traditional South Bohemian sounds). An essential part of all conferences is a social programme that always offers many opportunities to discuss various problems in an informal environment. The organizers have prepared a trip to the Kleť mountain including a pleasant guided tour of the Kleť Observatory and a conference dinner at the Kleť restaurant. Conference participants could hike to the top of the Kleť mountain or use the cableway.

The annual meeting of the CSOR decided that the 38th MME conference will be organized in the city of Brno by the Faculty of Business and Economics, Mendel University of Brno, on 9–11 September 2020.

13th Year of the *International Days of Statistics and Economics* (MSED 2019)

Tomáš Löster [1] | *University of Economics, Prague, Czech Republic*
Jakub Danko [2] | *University of Economics, Prague, Czech Republic*

From 5th to 7th September 2019, a worldwide conference of the International Days of Statistics and Economics (MSED) took place at the University of Economics in Prague. Every year, the conference enjoyed success and found participants from many countries around the world. The traditional goal of this international scientific conference was to present contributions by individual authors and discuss current issues in the areas of statistics, demography, economics and management and their interconnection.

The 13th International Days of Statistics and Economics was organized by The Department of Statistics and Probability and the Department of Managerial Economics, University of Economics, Prague, Czech Republic (main organizer); Faculty of Economics, The Technical University of Košice (co-organizer) and The Ton Duc Thung University, Ho Chi Minh City, Vietnam (co-organizer). The conference ranks among important statistical and economic conferences, which can be proved by the fact that Online Conference Proceedings from 2011 to 2018 were included in the Conference Proceedings Citation Index (CPCI), which has been integrated within the Web of Science, Clarivate Analytics (formerly Thomson Reuters). The output of the conference is on-line proceedings from the international scientific conference MSED published on the conference website. Presentations and subsequent discussions were held in English.

This year, 278 participants from various countries, such as the Russian Federation (87), Poland (37), Slovakia (16), registered at the conference. Other participants came from Vietnam, Austria, Bulgaria, France, Switzerland, etc. Doctoral students and young researchers from various universities abroad are also very happy to attend the conference. Among the participants, there were 35 doctoral students. Regarding statistical topics, the interest was traditionally focused on cluster analysis, computational statistics, and statistical modeling. Each contribution went through a double-blind peer-review process in accordance with publication ethics.

The key speaker during the conference was Helena Horská, the Head of Economic Research of the Raiffeisenbank, a.s., Czech Republic. Her lecture was accompanied by an interesting discussion about current economic topics. To conclude, we wish the conference be successful in the next year as well, because it is important that through this professional event deeper connections between important disciplines such as statistics and economics are established and the professional community realizes that mutual cooperation is crucial to the entire system.

We would also like to invite researchers, doctoral students and the wide professional public to the fourteenth International Days of Statistics and Economics, which will take place at the University of Economics, Prague, from 10th to 12th September 2020.

---

[1] Faculty of Informatics and Statistics, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. E-mail: tomas.loster@vse.cz.
[2] Faculty of Informatics and Statistics, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. E-mail: jakub.danko@vse.cz.
[3] More at: <*http://msed.vse.cz*>.

International Conference *Applications of Mathematics and Statistics in Economics* (AMSE 2019)

Stanislava Hronová [1] | *University of Economics, Prague, Czech Republic*

As usually, at the turn of August and September, from 28th August to 1st September 2019, already the 22nd international conference called *Applications of Mathematics and Statistics in Economics* took place. This year, the conference was organized by the Faculty of Economics of Matej Bel University in Banská Bystrica. As it is usual with Slovak organisers, the venue was in Slovak mountains, this time in *Nižná* in a picturesque foothills of the *Roháče* mountains. Over 60 experts from the Czech Republic, Slovakia, Poland, Austria, and the United Kingdom of Great Britain and Northern Ireland representing the University of Economics in Prague, Matej Bel University in Banská Bystrica, Wroclaw University of Economics, University of Economics in Bratislava, Vienna University of Economics and Business, Aston University in Birmingham, Comenius University in Bratislava, the Czech Statistical Office, and the Statistical Office of the Slovak Republic participated in the conference.

It is characteristic for the international conference that knowledge and experience are exchanged, the latest results of research are presented, and new procedures and methods are discussed there. Working meetings of representatives of cooperating workplaces and planning of further heading of scientific and pedagogical cooperation form an integral part of the conference. This year, the traditional meeting was supported by the auspices of presidents of the Czech Statistical Office and the Statistical Office of the Slovak Republic and by the participation of top representatives of both the offices.

It is because the conference was held in the year, in which we commemorate 100 years of Czechoslovak statistics; i.e. 100 years ago, the State Statistical Office in Prague was established. This anniversary, as well as the present and the future of the official statistics was reminded by representatives of both the statistical offices in their invited papers – Mr Marek Rojíček, President of the Czech Statistical Office, and Mr František Bernadič, Vice President of the Statistical Office of the Slovak Republic. Their papers provoked a rich discussion and interest of not only Czech and Slovak, but especially of foreign participants.

Among other important participants of the conference, who presented invited papers, were: Mikuláš Luptáčik, from the University of Economics and Business in Vienna, who in his paper *Efficiency vs. equity as a multi-objective optimisation problem* presented a new multiple criteria decision making model coupled with an extended Leontief input-output model taking into account the social dimension and obtain deeper insights into the so-called efficiency-equity trade-off, and Emmanuel Thanassoulis,

[1] Faculty of Informatics and Statistics, Department of Economic Statistics, University of Economics, Prague, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. E-mail: hronova@vse.cz.

from the Aston University in Birmingham, who in his paper called *The use of data envelopment analysis in regulation* showed how DEA has been integrated into regulatory systems and how the extension of DEA can improve the incentives aspect of regulation.

Other meetings of the conference took place in 9 sections as follows: Macroeconomic issues I, II, III, Social issues I, II, III, Multivariate statistical methods, Application in Insurance, and History of statistics. It is very difficult to highlight the most interesting papers. Therefore, I would like to underline only some papers of postgraduates and post-docs, which I consider to be of good quality, interesting, and innovative in terms of their methodology.

In the *Macroeconomic issues I* section, the most attractive was a paper called *How to value equity in National Accounts*, in which the authors (Kramulová, J., Vincenc, J., and Houžvičková, H.) presented possible methods of equity valuation with the aim to emphasize its relationship to revaluation of non-financial assets made in national accounts under the ESA 2010 manual. The paper highlights the worldwide initiative of the Czech Statistical Office in the field of equity valuation that results in the proposal of a new methodology of equity valuation, with an effort to stir up a debate about this topic. Their new methodology will undoubtedly be an inspiration also for other Member States of the EU.

In the *Macroeconomic issues II* section, papers based on utilization of an input-output analysis and DEA method were dominating. An exception to that was a paper of Gawthorpe, K. and Šafr, K., who in their paper called *Maintaining the well-being of ageing population in Czechia* presented government expenditures necessary to keep the current well-being unaltered to the dynamics of demographic prognosis. Their model is disaggregated with the so-called bottom-up approach to capture the nuanced differences among three selected ageing cohorts.

A practical application of multi-dimensional statistical methods (in the *Multivariate statistical methods* section) was presented by Stachová, M. and Král, P. in their paper called *Panel data clustering in financial distress prediction*. The goal of the paper has been to identify typical patterns in trajectories of financial indicators over time that could determine, via cluster analysis, whether a company tends to be in financial distress, or not, and also possibly the severity of this state.

In the *Applications in Insurance* section, what was undoubtedly interesting was the paper called *Equity release contracts with varying payments* (by the author: Marciuk, A.). The aim of her paper was to analyse varying payments of equity release contracts that have been in offer to clients in Poland for several years. Calculations were made on the basis of real Polish market data. Since the Solvency II directive requires the spot interest rate of the European Central Bank, the Svensson model was employed.

In the *Social issues III* section, the authors Čabla, A. and Habarta, F. in their theoretical paper called *Distribution of the wealth of the richest persons in the world* explored the probability distribution of wealth of the richest persons in the world based on estimates from the CEOWORLD magazine's rich list for March 2019.

A traditional section dealing with the *History of statistics* had a clearly determined topic this year, which was the 100th anniversary of establishment of the State Statistical Office in Prague. The paper called *Origin of the state statistical service in Czecho-Slovakia* (by the authors Závodský, P. and Šimpach, O.) was very interesting and it suitably completed the professional meeting of the conference the subtitle of which was namely to commemorate 100 years of the Czechoslovak official statistics.

A full programme of the AMSE 2019, including full texts of all presented papers can be found at: <*http://www.amse-conference.eu*>. There you can find also information about the history of the AMSE and links to preceding AMSE international conferences.

Papers presented at the AMSE 2019 conference are published in the book of proceedings that has been send to Thomson Reuters to be considered for inclusion into the Conference Proceedings Citation Index (CPCI). The proceedings of the past five AMSE conferences (i.e. AMSE 2014, 2015, 2016, 2017, and 2018) have been successfully indexed and are available in the Web of Science database.

The tradition of alternating organisation (Slovakia – Poland – the Czech Republic) further continues and the 23rd AMSE conference (to be organized by colleagues from the department of statistics of the Wroclaw University of Economics) will take place in *Wisla* in Poland at the turn of August and September 2020.

International Conference *Interdisciplinary Information Management Talks* (IDIMT 2019)

Petr Doucek [1] | *University of Economics, Prague, Czech Republic*
Lea Nedomová [2] | *University of Economics, Prague, Czech Republic*
Gerhard Chroust | *Johannes Kepler University Linz, Linz, Austria*
Antonín Pavlíček | *University of Economics, Prague, Czech Republic*

The Interdisciplinary Information Management Talks (IDIMT)[3] conference is among the conferences traditionally organized by the Department of Systems Analysis of the Faculty of Informatics and Statistics at the University of Economics in Prague, in cooperation with Johannes Kepler University Linz. Its predominant focus is on topics surrounding information management in various application fields, enterprise IT, and enterprise IT management. This year saw the conference's twenty-seventh iteration, bearing the subtitle "Innovation and Transformation in a Digital World." The conference's reputation is to be in close touch with current trends and a good indicator of future developments. The conference attracted papers from a total of 132 authors, with 43 submitted papers being accepted together with 11 invited papers. The authors have come from seven different countries: Austria, Czech Republic, Kazakhstan, Romania, Russia, Slovakia and Ukraine. It was held in early September in the historic town of Kutná Hora, in the pleasant environment of a former Jesuit campus that today houses the Central Bohemian Gallery (GASK).

The conference program was kicked off in the campus' beautifully ornamented refectory by Gerhard Chroust, professor emeritus at the University of Linz, who recalled the conference's long tradition and the changes over the twenty-five years of its existence. The next speaker at the gala plenary session was Jakub Fischer, dean of the Faculty of Informatics and Statistics, who welcomed the attendees and wished them a pleasant stay in Kutná Hora and great benefits from the conference, both from the conference papers themselves, and from informal meetings in the halls between presentations – meetings wherein new ideas and contacts for further research and teaching projects and collaboration often arise.

The next speaker, Christian W. Loesch – IBM's former director for Central and Eastern Europe – brought the conference from this general level down into individual topics, with a presentation entitled "ICT Future Scenarios: Visions and Challenges." His presentation interconnected the world of information and communication technologies with that of business and economics. Overall, it introduced some predicted scenarios for how the trends of artificial intelligence, quantum computing, supercomputing, and lateral ICT development might make a break-through into everyday economy. His presentation closed the introductory plenary block. It continued after a break with presentations in the conference's individual sessions.

---

[1] Faculty of Informatics and Statistics, University of Economics, Prague, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: doucek@vse.cz.

[2] Faculty of Informatics and Statistics, University of Economics, Prague, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: nedomova@vse.cz.

[3] More at: <*https://idimt.org*>.

Among the conference's most-visited session were "Digital Economy and Industry 4.0", "Innovations, New Business Models and Strategies", and "Social Media and On-Line Privacy." The conference's dominant focus will likely be clear from their names. The "Digital Economy and Industry 4.0" session presented the results of research on the digitalization of the economy in various countries and on the gradual transition towards a digital economy, primarily in industry, including the Internet of Things. Reflections on the development of the salaries of ICT employees formed another integral part of the papers. This session was well complemented by the visionary views in the "Society Beyond Industry 4.0: Smart Systems" session. The "Innovations, New Business Models and Strategies" session contained papers that presented specific impacts of information technologies on business models that are newly appearing in economics. This session also included papers on student start-ups and university activities aiming to support student business projects. The "Social Media and On-Line Privacy" session covered trending topics connected with the use of social media and its influence on human society. The remaining sessions primarily covered trends in information science. The "Digital Single Market Innovation" session covered the creation and presentation of a trust mechanism in the digital economic environment. The "Cybersecurity in a Digital World" session, by now a tradition at the conference, was dedicated to cybersecurity and information security – two highly relevant topics today. The theme of security appeared in other sessions as well, amounting to a thread that interwove the entire conference. The "Performance Management" session was a good complement to the conference's reflections on information science. It brought models for the deployment of new technologies back to the real world of costs and benefits.

This year, the conference included a half-day workshop dedicated to the EU Project "DRIVER+ (Driving Innovation in Crisis Management for European Resilience)". This project aims at supporting European Crisis Management (CM) by providing a large set of innovative solutions. Four papers presented relevant approaches: a digitalization of crisis management processes, an information strategy for CM, a dynamic state-of the-art catalogue of currently available crisis management solutions ('PoS'), and a method for evaluating the similarity of existing crisis management definitions in standards and guidelines. In the subsequent hands-on workshop participants learnt how to upload, search and download solutions in the PoS.

This conference was partially funded through project IGA 409039 of the Faculty of Informatics and Statistics.

Recent Publications

New publications of the Czech Statistical Office

Digitální ekonomika v číslech 2019 (Digital economy in numbers 2019) [online]. Prague: CZSO, 2019. <https://www.czso.cz/csu/czso/digitalni-ekonomika-v-cislech>.
Food Consumption in 2018. Prague: CZSO, 2019.
Generation, Recovery and Disposal of Waste for the period 2018. Prague: CZSO, 2019.
Využívání informačních a komunikačních technologií v domácnostech a mezi jednotlivci za období 2019 (Use of ICT in Households and by Individuals in 2019). Prague: CZSO, 2019.

Other selected publications

Ageing Europe. Looking at the lives of older people in the EU. 2019 Ed. Eurostat, 2019.
Energy, transport and environment statistics. 2019 Ed. Eurostat, 2019.

Recent Events

Conferences

The international scientific konference ***Quantitative Methods in Economics (QME 2020)***, organized by the Slovak Society for Operations Research and the Department of Operations Research and Econometrics, Faculty of Economic Informatics, University of Economics in Bratislava, will take place **from 27th to 29th May 2020 in Púchov, Slovakia**. More information available at: <*http://www.fhi.sk/ssov/conference*>.

The ***Robust 2020 Conference*** will be held **during 7–12 June 2020 in Bardejov, Slovakia**. More information available at: <*https://robust.nipax.cz*>.

The ***10th European Conference on Quality in Official Statistics (Q 2020)*** will be held **during 9–12 June 2020 in Budapest, Hungary**. More information available at: <*http://www.q2020.hu*>.

The ***23rd International Conference on Application of Mathematics and Statistics in Economics (AMSE 2020)***, organized by the Wrocław University of Economics, the University of Economics in Prague, and Matej Bel University in Banska Bystrica, will take place **from 26th to 30th August 2020 in Wisła, Poland**. More information available at: <*www.amse-conference.eu*>.

The ***Interdisciplinary Information Management Talks 2020 (IDIMT 2020)*** will take place **from 2nd to 4th September 2020 in Kutná Hora, Czech Republic**. More information available at: <*www.idimt.org*>.

The ***38th International Conference on Mathematical Methods in Economics 2020 (MME 2020)*** will be held **during 9–11 September 2020 in Brno, Czech Republic**. More information available at: <*https://mme2020.mendelu.cz*>.

The ***14th International Days of Statistics and Economics (MSED 2020)***, organized by the Department of Statistics and Probability and the Department of Managerial Economics of the University of Economics, Prague, Faculty of Economics of the Technical University of Košice and the Ton Duc Thang University, will be held **during 10–12 September 2020 in Prague, Czech Republic**. 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: <*https://msed.vse.cz*>.

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.

Discussion brings the opportunity to openly discuss the current or more general statistical or economic issues, in short what the authors would like to contribute to the scientific debate. Contribution shall have up to 6 000 words or up to 10 1.5-spaced pages.

In 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.

Consultation contains papers focused primarily on new perspectives or innovative approaches in statistics or economics about which the authors would like to inform the professional public. Consultation shall have up to 6 000 words or up to 10 1.5-spaced pages.

The *Book Review* section brings reviews of recent books in the fieled of the official statistics. Reviews shall have up to 600 words or one (1) 1.5-spaced page.

The *Information* section includes informative (descriptive) texts, information on latest publications (issued not only by the CZSO), recent and upcoming scientific conferences. Recommended 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.

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.* 1st Ed. 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. AND 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: Statistics and Economy 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 and numbered.

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. Articles for the review process are accepted continuously and may contain tables and figures in the text (for final graphical typesetting must be supplied separately as specified in the instructions above). Please be informed about our Publication Ethics rules (i.e. authors responsibilities) published at: <http://www.czso.cz/statistika_journal>.

Managing Editor: Jiří Novotný
phone: (+420) 274 054 299
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 117 or USD 174 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: Václav Adam
Print: Czech Statistical Office

100th year of the series of professional statistics and economy journals of the State Statistical Service in the Czech Republic: ***Statistika*** (since 1964), ***Statistika a kontrola*** (1962–1963), ***Statistický obzor*** (1931–1961) and ***Česko-slovenský statistický věstník*** (1920–1930).

Published by the Czech Statistical Office
ISSN 1804-8765 (Online)
ISSN 0322-788X (Print)
Reg. MK CR E 4684