Ministry of Finance
of the Czech Republic

Forecast of Public Health Insurance Revenues

Aleš Bělohradský and Zdeněk Štolc

Methodological Compendium

Forecast of Public Health Insurance Revenues
Aleš Bělohradský and Zdeněk Štolc

Methodological Compendium

Ministry of Finance of the Czech Republic
Letenská 15, 118 10 Prague 1

First edition, Praha 2018

Email: Ales.Belohradsky@mfcr.cz, Zdenek.Stolc@mfcr.cz

ISBN 978-80-7586-023-1 (on-line)

Electronic archive:
http://www.mfcr.cz/methodology

Forecast of Public Health Insurance Revenues

Aleš Bělohradský and Zdeněk Štolc

Methodological Compendium

Contents

The Methodological Compendium presents the methodological approaches of the Ministry of Finance of the Czech Republic in the areas of macroeconomic and fiscal analyses, forecasts and projections.

We will gladly welcome relevant comments or suggestions to improve the publication. Please send any comments to the author of the publication.

Introduction and Summary

Forecasting public health insurance revenues is essential not only for health insurance companies, but also for the Ministry of Finance, for example when setting medium-term expenditure frameworks of the state budget and state funds within the budgetary process. The revenue forecast is used by the health insurance companies, because they are obliged to prepare an annual health insurance plan according to Act No. 551/1991 Coll. and Act No. 280/1992 Coll. Then Act No. 23/2017 Coll. requires health insurance companies to create a medium-term outlook for revenues and expenditures for the next two years.

In connection with the Fiscal Responsibility Rules Act, the Ministry of Finance draws up a binding government document for the preparation of the draft state budget, the budgets of state funds and their medium-term outlook – the Budgetary Strategy of the General Government Sector of the Czech Republic. The document contains expenditure frameworks of the state budget and state funds or financial relations of the state budget to the special account of public health insurance. As part of establishing medium-term expenditure frameworks, the forecast of public health insurance revenues represents both a component of total general government revenues, from which total general government expenditure is derived, as well as one of the key variables for determining the structural balance of the general government sector components outside the state budget and state funds.

To forecast revenues from public health insurance, it is crucial to determine the amount of revenues from health insurance contributions, since it accounts for about 99% of the total revenues of health insurance companies. The other components of the revenues of health insurance companies can be considered negligible. The amount of the health insurance contribution is generally determined by the assessment base, the number of insured persons and the health insurance contribution rate (for all payer groups it is 13.5% of the relevant assessment base). Since Act No. 592/1992 Coll. defines different assessment bases for different insurance payers, the determination in this methodology of the public health insurance revenues is based on the specification of partial models for the forecast of insurance health insurance contributions from individual payers. The total public health insurance revenues are then determined by aggregating the partial forecasts of health insurance contributions of the relevant payers.

To formulate the partial models, use is made of quarterly time series of relevant variables for the years 2000–2018. In all time series, a logarithmic transformation is used to reduce the level difference and the difference in the variance of the individual series. Because of the non-stationarity of time series, the variables are expressed in the first differences.[1] The forecast accuracy of the models is assessed for each specification based on the average root mean square error.[2] According to the minimum value of this criterion, use was made of the estimate using the modified least squares method as the most appropriate estimate to forecast health insurance contributions from employees; for health insurance contributions paid by the state for the state insured persons, the regression model includes, as an independent variable, the assessment base and the number of unemployed registered with the Labour Office. In the case of self-employed persons, the error correction model appears to be the best of the tested models, and for persons without taxable incomes it is a simple process of first order moving averages without the logarithmic transformation of variables. As a reference model, a “naïve” first order autoregression process is then chosen for all payers without further independent variables.

In the period between 12 July and 17 August 2018, the material was subject to consultation with the Ministry of Health, the General Health Insurance Company and the Association of Health Insurance Companies. Their relevant comments were subsequently incorporated in the methodology. Commenting institutions expressed their agreement with the concept of partial forecasts according to individual payers as well as with selected model specifications, including selected independent variables. The methodology has been found to be useful for forecasting the revenues of the entire public health insurance system.

The methodology of the forecast of public health insurance in the Czech Republic is divided into six parts. The following chapter deals with a description of the revenue structure of the public health insurance system. The next four chapters are then broken down by individual payers and contain the specification of suitable forecast models. In conclusion, we present the overall result of the forecast model.

[1] All variables used in the following models are non-stationary, which we consider to be a common feature of macroeconomic data. They do not respond to deterministic processes, which is tested by the augmented Dickey-Fuller test and the Phillips-Perron test. The first difference is therefore a natural choice. In the case of employees, we use the year-on-year difference due to the significantly seasonal nature of the time series; in other cases (where there are often surges in the assessment bases) a quarterly difference is used. The logarithmic transformation is used due to the different scales of the series. Transformed time series are stationary unless otherwise specified in the text, which is always tested by the augmented Dickey-Fuller test, although the results of this test are not presented in the text.

[2] This is the average for the last 4 years, when the values were derived from a series of 16 *in-sample* forecasts for moving two-year periods. The advantage of this approach is the elimination of one-off fluctuations. Considering the *in-sample* forecast only for the last period (which we use for graphical depiction) could lead to the choice of a model that corresponds to one period but deviates in the long run.

1 Financing of Health Care in the Czech Republic

The System of Health Accounts 2011 (OECD, 2017) distinguishes three schemes of healthcare financing: public revenue (i.e., compulsory health insurance contributions and public budget contributions, including state and local budgets), household out-of-pocket payments and voluntary payments other than household out-of-pocket payments. In 2016, total healthcare expenditure in the Czech Republic amounted to CZK 361.6 billion, or 7.6% of GDP. This expenditure is mainly covered by health insurance companies from public health insurance (about 66% of total expenditure), in some cases the reimbursement is supplemented by the co-payment of patients (e.g. co-payment for some medicines, dental care, etc.). Specific activities not covered by the public health insurance (e.g. medical research expenditure) are financed from state and local budgets.

Table 1: Health-care Expenditure in the Czech Republic by Financing Scheme

		2010	2011	2012	2013	2014	2015	2016
Total health-care expenditure	*CZK bn.*	**334.1**	**338.1**	**341.9**	**344.7**	**348.7**	**352.0**	**361.6**
	% of GDP	**8.4**	**8.4**	**8.4**	**8.4**	**8.1**	**7.7**	**7.6**
Government and compulsory contributory health	*CZK bn.*	**282.2**	**284.8**	**288.6**	**292.1**	**291.6**	**293.4**	**300.2**
care scheme	*% of GDP*	**7.1**	**7.1**	**7.1**	**7.1**	**6.8**	**6.4**	**6.3**
Compulsory contrib. health insurance scheme	*CZK bn.*	231.9	234.3	237.9	238.4	234.6	234.5	237.7
Central government scheme	*CZK bn.*	45.2	44.6	44.6	47.5	50.7	52.6	55.9
Local government scheme	*CZK bn.*	5,1	5.9	6.0	6.2	6.3	6.3	6.6
Voluntary health-care payment scheme	*CZK bn.*	**9.2**	**9.2**	**9.1**	**9.2**	**10.7**	**9.3**	**10.2**
	% of GDP	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**
Voluntary health insurance scheme	*CZK bn.*	0.4	0.5	0.5	0.5	0.5	0.5	0.5
Non-profit institutions financing scheme	*CZK bn.*	7.9	7.8	7.7	7.7	7.8	7.9	8.5
Enterprise financing scheme	*CZK bn.*	0.9	0.9	0.9	1.0	2.4	0.9	1.3
Household out-of-pocket payment	*CZK bn.*	**42.7**	**44.0**	**44.2**	**43.5**	**46.5**	**49.4**	**51.2**
	% of GDP	**1.1**	**1.1**	**1.1**	**1.1**	**1.1**	**1.1**	**1.1**

Note: Figures are expressed in accrual basis based on the System of Health Accounts 2011. According to the methodology payments for state insured persons from the state budget are allocated to the revenues of health insurance companies. Data for household out-of-pocket payments are preliminary in 2016.
Source: CZSO (2018a).

The healthcare system in the Czech Republic is characterized by the fact that health insurance companies, whose revenues come from public health insurance, have a key position in terms of financing health care. In accordance with Section 2(1) of Act No. 48/1997 Coll., every person with permanent residence on the territory of the Czech Republic must be covered by public health insurance.[3] In accordance with Section 4 of this Act, public health insurance contribution payers are employees, employers, self-employed persons and persons without taxable incomes. Payer is also the state which, through the state budget, pays the health insurance contribution for the "state insured persons".[4] Public health insurance contribution thus finances all health care guaranteed by law.

Table 2: Revenue Structure of the Public Health Insurance System

		2011	2012	2013	2014	2015	2016	2017
Total revenue	*CZK bn.*	**220.1**	**223.6**	**228.6**	**241.3**	**252.6**	**264.9**	**284.8**
Health insurance contributions	*CZK bn.*	**217.4**	**221.0**	**225.8**	**238.5**	**249.7**	**261.8**	**281.6**
Employers	*CZK bn.*	148.0	151.4	155.2	161.4	170.5	180.2	195.8
	% of health ins. contribution	68.1	68.5	68.8	67.7	68.3	68.8	69.5
Self-employed persons	*CZK bn.*	14.1	14.2	14.3	14.7	15.5	16.0	16.9
	% of health ins. contribution	6.5	6.4	6.3	6.2	6.2	6.1	6.0
Persons without taxable income	*CZK bn.*	2.6	2.5	2.5	2.5	2.8	3.3	3.6
	% of health ins. contribution	1.2	1.1	1.1	1.1	1.1	1.3	1.3
State insured persons	*CZK bn.*	52.7	52.9	53.7	59.9	60.9	62.3	65.3
	% of health ins. contribution	24.3	23.9	23.8	25.1	24.4	23.8	23.2
Other revenue	*CZK bn.*	**2.7**	**2.6**	**2.8**	**2.8**	**2.9**	**3.1**	**3.2**

Note: Figures are expressed in cash basis.
Source: Annual Reports of health insurance companies. Ministry of Finance calculations.

[3] Also, persons who do not have permanent residence in the Czech Republic if they are employees of an employer with a registered office or permanent residence in the Czech Republic.
[4] State insured persons are defined exhaustively in Section 7(1) of Act No. 48/1997 Coll.

The largest part of the revenues of health insurance companies (about 99%) consists of selected health insurance contributions from the mentioned groups of payers. Revenues from sanctions, penalties or income from foreign insurance companies play a marginal role in the revenues of health insurers.

Considering the revenue structure of the public health insurance system, it is clear that the health insurance contributions are crucial to forecast the revenues of the system, and other revenue components of health insurers can be considered negligible. As a result of the different determination of the assessment base for the health insurance contributions payable by individual payers (Table 3), forecast models for individual groups of payers are compiled.

Table 3: Assessment Base for the Health Insurance Contribution Payers

	Assessment base	Minimum assessment base	Maximum assessment base
Employees	total income	minimum wage	In 2008–2009 specified as 48 multiple of average wage; in 2010–2012 as 72 multiple of average wage; since 2013 there is no force
Self-employed persons	50% of profit	12 multiple of 50% product of general assessment base for pension insurance in given year (2 years preceding) and conversion coefficient	In 2008–2009 specified as 48 multiple of average wage; in 2010–2012 as 72 multiple of average wage; since 2013 there is no force
Persons without taxable income	minimum wage	×	×
State insured persons	resolution of the government	×	×

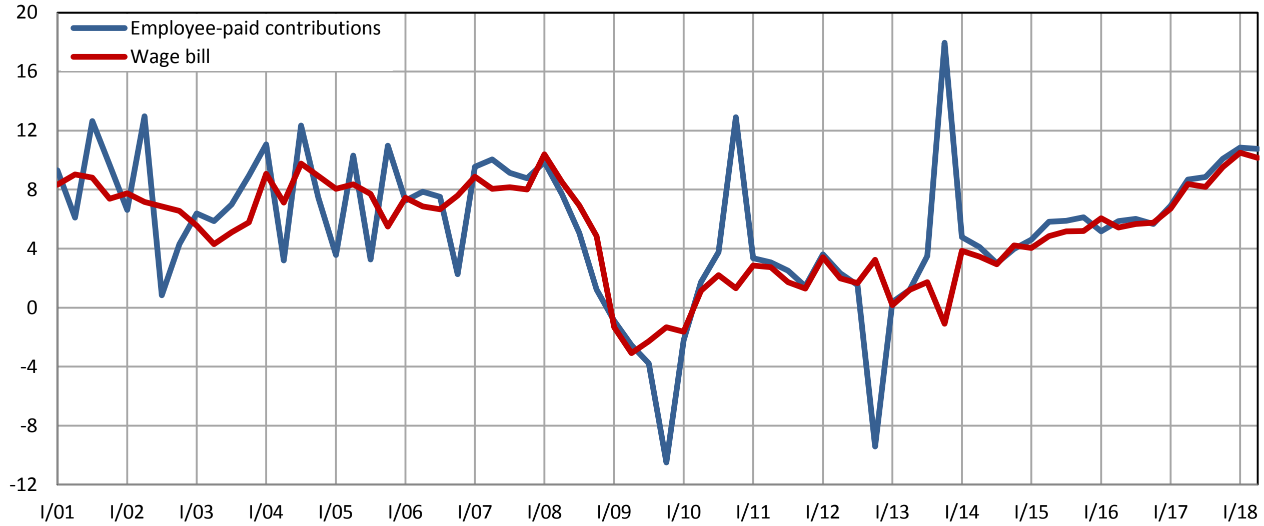
Source: Act No. 592/1992 Coll. Adjustment by Ministry of Finance.

2 Employees

With regard to the construction of assessment base for the employee-paid health insurance contributions (Table 3), the wage bill is selected as the core independent variable. Alternatively, the separate inclusion of the number of employees and the average wage was also tested, but the differentiation had little effect on the result.

Figure 1: Revenues from Employee-paid Contributions and Wage Bill

(Year-on-year growth rate in %)



Source: Annual Reports of health insurance companies, Act No. 592/1992 Coll.; CZSO (2018b).

Figure 1 shows how closely the revenue from employee health insurance contributions reflects the development of wage bill. However, we can find periods in which the series behave differently. Until 2008, health insurance contribution collection was significantly more volatile than wages in comparison to the following period. In that period, there were two quarters showing significant downward movements (mirrored in annual differences the quarter a year later), namely Q4 of 2009 and 2012. In these periods, there were non-systematic deviations, probably linked to the changes in the statutory ceiling for general health insurance, as outlined in Table 3. To avoid impact of the deviation on the resulting estimates, the dummy variables are included in the model in all specifications. To take into account any variability in quarterly dynamics, we also use seasonal dummy variables.

The basic specification of the employee health insurance contribution collection model uses the ordinary least squares (OLS) estimate.

$$\Delta\log(HIC\_EMP_t) = \alpha_0 + \alpha_1 \Delta\log(WB_t) + \delta_s Q_s + \varepsilon_t \tag{1}$$

where α_0, α_1 are the model parameters, $\Delta(\bullet)$ is the year-on-year difference of the given variable, HIC_EMP_t is the collected employee health insurance contribution at time t, WB_t is the wage bill at time t, Q_s is a seasonal dummy variable with parameters δ_s, and ε_t indicates an error term.

Based on Johansen test, cointegration of first order was confirmed between the employee-paid health insurance contributions and the wage bill. The year-on-year differences in the model (1) do not clearly guarantee stationary time series, although this depends on the test. Therefore, we also consider the Error Correction Model (ECM)

$$\Delta\log(HIC\_EMP_t) = \alpha_0 + \alpha_1 \Delta\log(WB_t) + \alpha_2 \hat{u}_{t-1} + \delta_s Q_s + \varepsilon_t \tag{2}$$

where $\hat{u}_{t-1}$ are the residuals at time *t-1* obtained from the regression model

$$\log(HIC\_EMP_t) = \beta_0 + \beta_1 \log(WB_t) + u_t \tag{3}$$

As an alternative way of dealing with cointegration, DOLS (Dynamic OLS) and FMOLS (Fully Modified OLS) estimation methods were also tested. DOLS includes a long-term relationship between variables by adding lagging and future values, while FMOLS modulates the covariance matrix for regression coefficients estimation taking into account long run relationships. Although the time series of employee health insurance contributions tend to be rather stochastic, a specification with linear trend in equation (3) has also been estimated, resulting in statistical improvement of the model. Table 4 provides the results for selected model specifications and estimation methods.

Due to higher volatility in the first sample period[5] (see Figure 1), it was also envisaged to reduce the estimated period to the last 10 and 5 years, respectively. In both cases, however, the relevant in-sample forecast showed a significantly higher average root mean square error (RMSE). Table 4 therefore shows only estimates for the whole period. In addition to the reference AR(1) process, another robustness check has been tested: direct dependence of the employee-paid health insurance contribution on the wage bill development, i.e. that the two variables would develop identically.

Table 4: Regression Results for Revenues from Employee-paid Contributions

Employee-paid contributions	OLS	DOLS	FMOLS	ECM c	ECM d	WB	naïve AR(1)
Constant	0.00	0.01	0.00	-0,01	0.00		0.03 ***
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)		(0.01)
Wage bill	0.95 ***	0.92 ***	0.98 ***	1.02 ***	0.95 ***	1.00	
	(0.08)	(0.09)	(0.05)	(0.08)	(0.07)		
AR(1)							0.50 ***
							(0.11)
Error correction				-0.48 ***	-0.55 ***		
				(0.12)	(0.10)		
Number of observations	70	66	69	70	70		69
Adj. R-squared	0.80	0.80	0.80	0.84	0.86		0.23
AIC	-4.78			-4.99	-5.13		-3.50
Average RMSE	0.30	0.35	0.29	1.12	0.50	0.34	1.30

*Note: standard errors in parentheses, * p < 0.10; ** p < 0.05; *** p < 0.01. OLS – equation (1), all models comprise seasonal dummy variables, which are, however, close to zero, thus, they are omitted from the table; ECM denotes error correction model, while c is without trend and d including linear trend in the cointegration equation (labels c and d correspond to the standard classification of ECM); column WB represents direct dependence on the wage bill, i.e. it is not a regression model; naïve AR(1) process is the reference model. For the assessment of the models, we use the adjusted R-squared, Akaike information criterion (AIC) and average root mean square errors (RMSE) computed from two-years moving span (in-sample) forecasts within last 4 years.*

Source: Ministry of Finance calculations.

According to the average RMSE, the lowest forecast error, despite the reported volatility in the first years, is in the estimate for the entire sample. The estimates of OLS, DOLS and FMOLS are not much different, yet FMOLS has a slightly lower RMSE and, moreover, it deals with the non-stationarity of year-to-year differentiated logarithmically transformed data. Only a slightly higher error rate is shown by the non-regression model for direct linking to the development of wage bill, especially in the medium term. For this reason, the resulting model is the combination of the FMOLS model, which also envisages quarter-on-quarter variability, for the first two forecast years, and a simple connection to the wage bill for the subsequent years of outlook.

$$\Delta\log(HIC\_EMP_t) = \begin{cases} \gamma_0 + \gamma_1 \Delta\log(WB_t) + \delta_s Q_s + \varepsilon_t & t \in [1, T-12] \\ \left(\frac{T-t}{4} - 2\right) \cdot \left(\gamma_0 + \gamma_1 \Delta\log(WB_t) + \delta_s Q_s + \varepsilon_t\right) + \left(3 - \frac{T-t}{4}\right) \cdot \Delta\log(WB_t) & t \in [T-11, T-8] \\ \Delta\log(WB_t) & t \in [T-7, T] \end{cases} \quad (4)$$

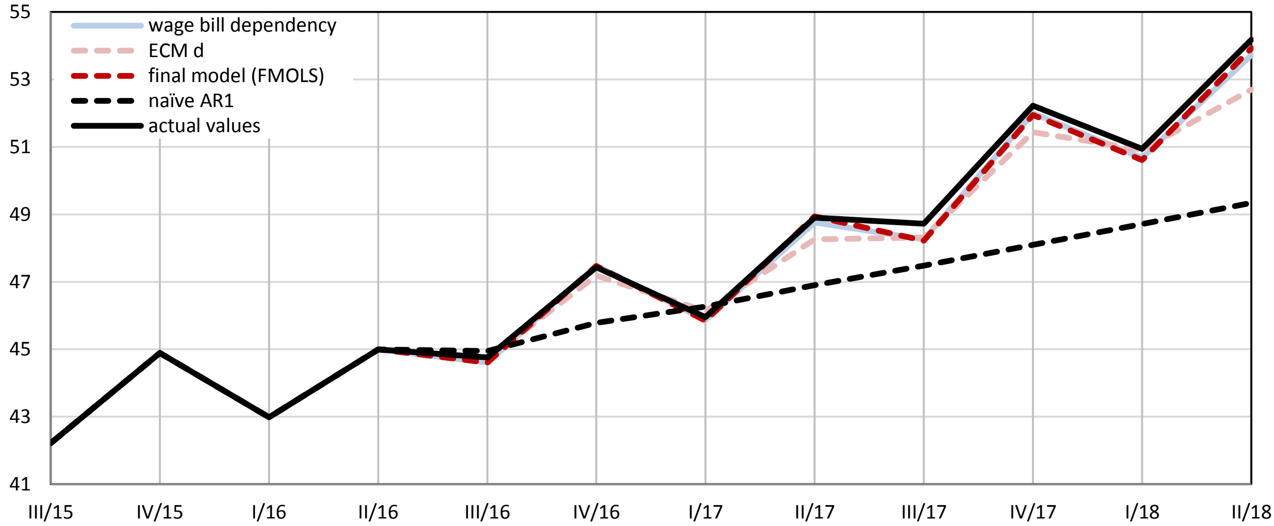
where γ_0, γ_1 indicate the model parameters calculated by the FMOLS method. Therefore, this differs from the estimates of α_0, α_1 calculated by the basic OLS method. $\Delta(\bullet)$ points to the year-on-year difference of the given variable, HIC_EMP_t is selected health insurance contributions for employees, WB_t is the wage bill, Q_s represents seasonal dummy variable, ε_t is an error term and T indicates the horizon of the outlook in the number of quarters. For the last two years, employee-paid contribution growth is linked directly to the wage growth. In the previous four quarters, the growth rate of employee health insurance contributions is linearly approaching the growth rate of the wage bill.

The resulting model, which in the medium term is identical to the wage growth, reduces the average RMSE slightly to 0.287 in the ten most recent predictions. Figure 2 shows the resulting model, not the FMOLS itself (the difference would not be visible on this scale).

[5] Tested using the Goldfeld-Quandt test with time division of data until 2006 and after 2007.

Figure 2: Comparison of Actual and Forecasted Revenues from Employee-paid Contributions

(Billions of CZK)



Note: the final model is constructed as a combination of FMOLS and direct linkage to the wage bill.

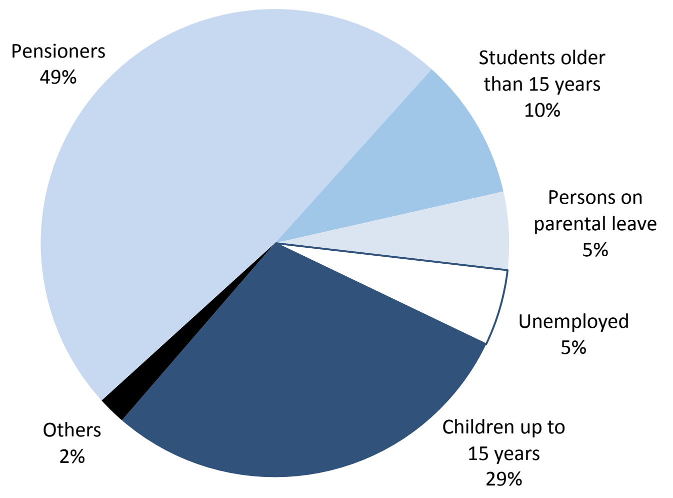
Source: Annual Reports of health insurance companies, Ministry of Finance calculations.

3 State Insured Persons

The group of persons whose health insurance contributions are paid by the state (i.e. state insured persons) is defined in Section 7(1) of the Act No. 48/1997 Coll. It is clear from the structure of the state insured persons (Figure 3) that the number of pensioners, children, students and the number of unemployed people is decisive. The amount of health insurance contributions is influenced primarily by the assessment base for the state insured persons under Section 3c of Act No. 592/1992 Coll. and in the relevant Government Regulations.

Figure 3: Structure of State Insured Persons (2017)

(In % from all state insured persons)

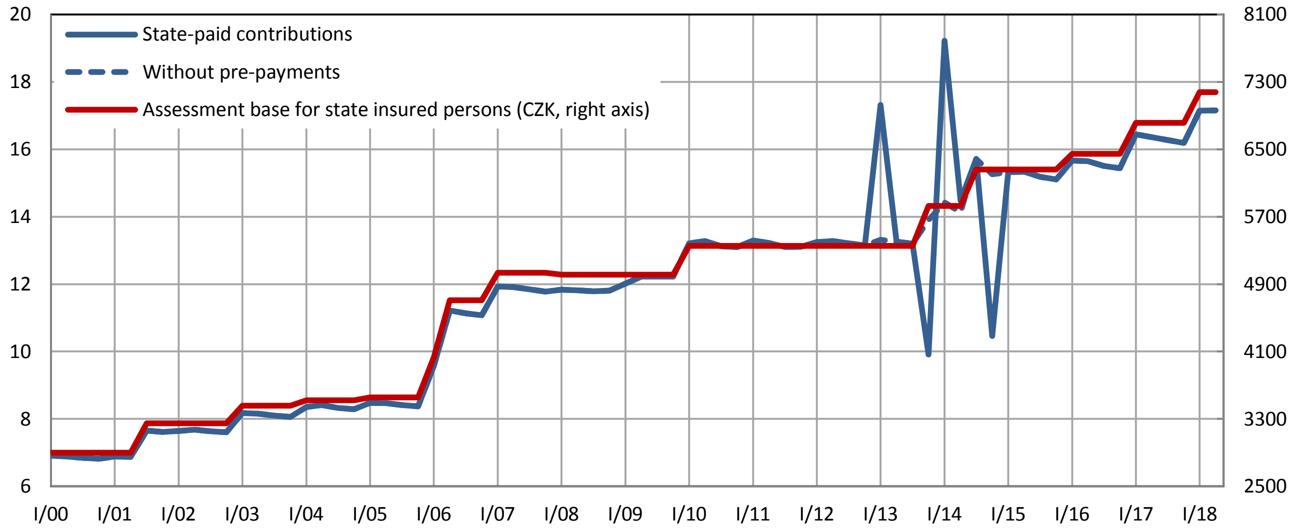


Source: Ministry of Finance.

Figure 4 illustrates the close relationship between the assessment base and the health insurance contributions collected for the state insured persons. Two significant fluctuations between the first and the last quarter of 2013 and 2014 are so-called "pre-payments", essentially interest-free loans from the state budget to health insurance companies for the purpose of short-term stabilization of their economy.[6]

Figure 4: Revenues from State-paid Contributions and Assessment Base for State Insured Persons

(Billions of CZK on left axis, CZK on right axis)



Source: Annual Reports of health insurance companies, Act No. 592/1992 Coll. and relevant Government Regulations; adjusted by the Ministry of Finance.

[6] In the first quarter of 2013, a prepayment of CZK 4 billion was provided, and repaid in Q4 2013. In the first quarter of 2014, the second prepayment of CZK 4.8 billion was provided, and repaid in Q4 2014.

The basic specification of the model is therefore as follows:

$$\Delta\log(HIC\_SIP_t) = \alpha_0 + \alpha_1\Delta\log(AB_t) + \alpha_2\Delta\log(OA\_PENS_t) + \alpha_3\Delta\log(U_t) + \delta_s Q_s + \varepsilon_t \quad (5)$$

where $\alpha_0, \dots, \alpha_3$ indicate the model parameters, $\Delta(\bullet)$ indicates the first difference of the variable, HIC_SIP_t means health insurance contributions paid by the state for the state insured persons at time t less pre-payments, AB_t is the assessment base determined by the relevant Government Regulations, OA_PENS_t is the number of old-age pensioners at time t, U_t indicates the number of registered unemployed persons according to the statistics of the Ministry of Labour and Social Affairs, Q_s is a seasonal dummy variable with parameters δ_s (as in the case of employees) and ε_t indicates an error term. In alternative specifications, we also tested the number of children and students, but it did not provide any explanatory information.

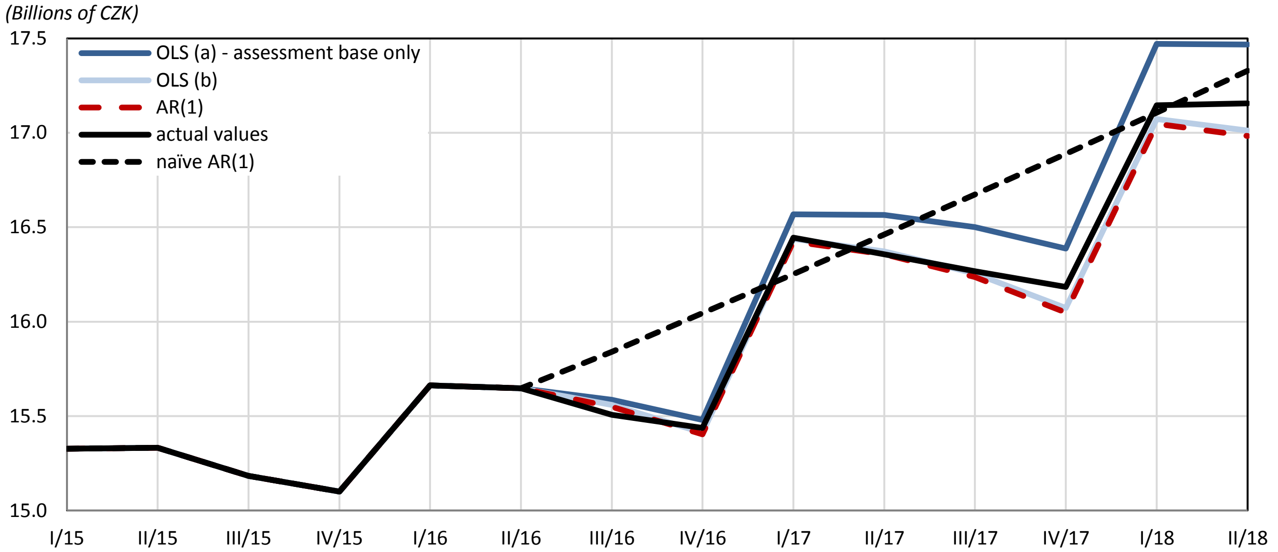
Model (5), however, contains autocorrelation, which is removed, based on the respective tests, by the first-order autoregression model. Table 5 provides the results of these two specifications, complemented by the model including only the assessment base and the reference AR(1) process.

Table 5: Regression Results for Revenues from State-paid Contributions

State-paid contributions	OLS (a)	OLS (b)	AR(1)	naïve AR(1)
Constant	-0.01 ***	-0.01 ***	-0.01	0.01 **
	(0.00)	(0.00)	(0.00)	(0.00)
Assessment base for state insured persons	0.99 ***	0.99 ***	1.00 ***	
	(0.03)	(0.03)	(0.03)	
Number of old age pensioners		0.06		
		(0.15)		
Unemployment		0.05 **	0.05 ***	
		(0.02)	(0.01)	
AR(1)			-0.63 ***	0.02
			(0.10)	(0.12)
Number of observations	73	73	73	72
Adj. R-squared	0.94	0.94	0.97	-0.01
AIC	-6.61	-6.62	-7.11	-3.80
Average RMSE	0.27	0.21	0.20	0.71

*Note: standard errors in parentheses, * p < 0.10; ** p < 0.05; *** p < 0.01. OLS (a) – equation (5) only with assessment base as an explanatory variable; OLS (b) – equation (5); AR(1) – equation (6) including AR (1) process using maximum likelihood estimation (not OLS); AR – reference AR (1) process. All models except the naïve AR(1) comprise also seasonal dummy variables, which have positive and statistically significant coefficients, however, relatively close to zero, thus, they are omitted from the table.*
Source: Ministry of Finance calculations.

Figure 5: Comparison of Actual and Forecasted Revenues from State-paid Contributions
(Billions of CZK)



Source: Annual Reports of health insurance companies, Ministry of Finance calculations.

Comparison of actual values with in-sample predictions shows that the regression model including only the assessment base tends to overstate the forecast. The inclusion of unemployment (which has fallen sharply in recent years) is therefore important. The effect of the lagged dependent variable is statistically significant. According to the average for the last 16 quarters, the autoregressive model offers lower forecast error, thus, the effect of the increase of the assessment base for the state insured persons is highest in the first quarter after the increase and then gradually decreases. We therefore include the autoregressive element in the resulting specification.

$$\Delta \log(HIC\_SIP_t) = \alpha_0 + \alpha_1 \Delta \log(AB_t) + \alpha_2 \Delta \log(U_t) + \alpha_3 \Delta \log(HIC\_SIP_{t-1}) + \delta_s Q_s + \varepsilon_t \quad (6)$$

Coefficients $\alpha_0, \dots, \alpha_3$ and δ_s are estimated using the maximum likelihood method.

4 Self-employed Persons

According to Table 3, the minimum assessment base may be considered as an independent variable for self-employed persons. The minimum assessment base is provided by Section 3a(2) of the Act No. 592/1992 Coll. Next, we assume the influence of the number of self-employed persons and also the impact of the economic situation in the country affecting the profitability, which we include through nominal GDP. We choose the nominal GDP because the health insurance contributions and other variables are expressed in nominal terms as well.

While in previous groups of payers was a clear correlation with one of the independent variables, the situation is less clear for the self-employed persons. The development in health insurance contributions from self-employed persons can roughly be compared to the development of nominal GDP and the minimum assessment base; however, it is significantly more volatile and does not show a distinct seasonal character as in the case of the independent variable. For this reason, a better fitting estimation may be provided by the smoothed time series using the Hodrick-Prescott (HP) filter, or for the annual time series (see below). The relatively linear trend of the time series for health insurance contributions may also cause inconsistency in the case of stacionarization by the first difference (as it is not invertible). The presence of autocorrelation, however, corresponds rather to the moving average (MA) process. Moreover, all three variables are cointegrated, so an error correction model is also used.

Table 6: Regression Results for Revenues from Self-employed Payers – Quarterly Data

Contributions from self-employed persons	OLS	MA(1)	ECM c	naïve AR(1)	HP filter-OLS	HP filter-ECM
Constant	0.05 *	0.02 *	-0.00	0.02	0.01 ***	0.01 ***
	(0.03)	(0.01)	(0.02)	(0.02)	(0.00)	(0.00)
Minimum assessment base	-1.83		1.38		0.04	0.23 ***
	(1.18)		(0.97)		(0.04)	(0.07)
Nominal GDP	-1.31 ***	-0.28	-0.06		0.24 ***	0.09 ***
(seasonal adjusted in the case of HP filter)	(0.46)	(0.52)	(0.36)		(0.08)	(0.03)
Number of self-employed persons	0.89				0.13 *	0.01
	(1.35)				(0.07)	(0.07)
AR(1)				-0.48 ***		
				(0.10)		
MA(1)		-0.74 ***				
		(0.14)				
Error correction			-0.96 ***			-0.09 ***
			(0.13)			(0.03)
Number of observations	72	72	73	72	73	73
Adj. R-squared	0.08	0.37	0.48	0.22	0.11	0.16
AIC	-0.81	-1.17	-1.38	-1.01	-6.63	-6.67
Average RMSE	0.55	0.27	0.16	0.24	0.22	0.22

*Note: standard errors in parentheses, * p < 0.10; ** p < 0.05; *** p < 0.01. Adjusted R-squared and AIC for models with HP filter are not comparable with the others, because they correspond to the smoothed series. However, RMSE is comparable.*
Source: Ministry of Finance calculations.

A simple estimate for the unsmoothed series using OLS (Table 6) shows a minimal explanatory capability and a significant forecast error. Although the error correction model has the best explanatory capability, the correction term is the only statistically significant variable (coefficient close to -1 means that the contributions exactly respond to the previous year's deviation from the value corresponding to the number of self-employed persons and quarterly GDP according to equation (8)). Due to the strong and hard-to-predict instability of the dependent variable, it seems more reasonable to predict smoothed values, i.e. regardless of short-term deviations. However, it is not stationary even in log-differences. As it is cointegrated with seasonally adjusted nominal GDP, the error correction model (ECM) is again the most appropriate specification.

$$\Delta \log[HP(HIC\_SEP_t)] = \alpha_0 + \alpha_1 \Delta \log(MAB_t) + \alpha_2 \Delta \log(GDP\_SA_t) + \alpha_3 \Delta \log(SEP_t) + \alpha_4 \hat{u}_{t-1} + \varepsilon_t \quad (7)$$

where $HP(\bullet)$ indicates the use of the Hodrick-Prescott filter (λ = 100), $\alpha_0, \dots, \alpha_4$ are the model parameters, $\Delta(\bullet)$ is the first difference of a given variable, HIC_SEP_t indicates collected health insurance contributions from the self-employed persons at time t, MAB_t indicates the minimum assessment base for the payment of health insurance contributions by self-employed persons at time t, GDP_SA_t indicates seasonally adjusted nominal GDP at time t, SEP_t

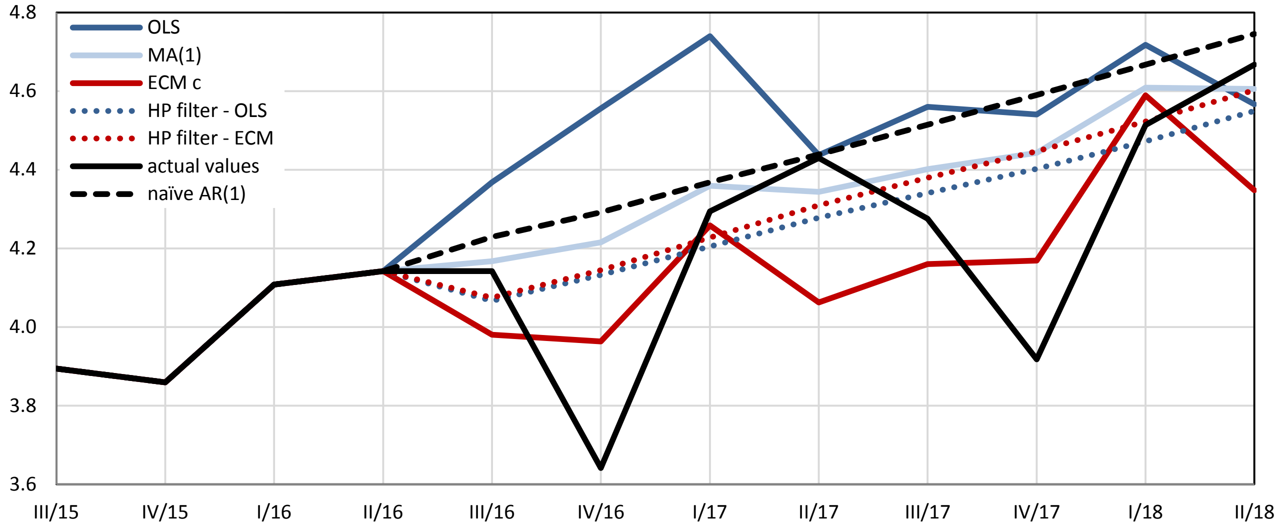
represents the number of registered self-employed persons, ε_t is an error term and $\hat{u}_t$ are the residuals obtained from the regression model

$$\log(HIC\_SEP_t) = \beta_0 + \beta_1 \log(MAB_t) + \beta_2 \log(GDP\_SA_t) + \beta_3 \log(SEP_t) + u_t \quad (8)$$

The model (7) shows slightly higher errors (see the last column of Table 6) than the error correction model for the smoothed series.

Figure 6: Comparison of Actual and Forecasted Revenues from Self-employed Payers – Quarterly Data

(Billions of CZK)



Source: Annual Reports of health insurance companies, Ministry of Finance calculations.

Apart from the HP filter, in order to overcome the volatility of quarterly data on self-employed persons' contributions, we also run the estimation on annual data. All in all, we ultimately focus on the annual forecasts and, if quarterly data fail to explain the quarter-on-quarter variability, annual data can bring more accurate results. We keep the same specifications as for quarterly data. Even in this case, the first order moving average process and cointegration between variables are present.

Table 7: Regression Results for Revenues from Self-employed Payers – Annual Data

Contributions from self-employed persons	OLS	MA(1)	ECM c	naïve AR(1)	ECM (Q)	HP ECM (Q)
Constant	-0.03		-0.01	0.05 **	-0.00	0.01 ***
	(0.03)		(0.01)	(0.02)	(0.02)	(0.00)
Minimum assessment base	0.53		0.18		1.38	0.23 ***
	(0.49)		(0.24)		(0.97)	(0.07)
Nominal GDP	1.15 **	1.00 ***	1.28 ***		-0.06	0.09 ***
	(0.44)	(0.18)	(0.22)		(0.36)	(0.03)
Number of self-employed persons	0.69	1.08 **	0.68 **			0.01
	(0.49)	(0.49)	(0.24)			(0.07)
AR(1)				0.06		
				(0.23)		
MA(1)		-0.59				
		(0.41)				
Error correction			-1.18 ***		-0.96 ***	-0.09 ***
			(0.18)		(0.13)	(0.03)
Number of observations	17	17	17	16	73	73
Adj. R-squared	0.33	0.42	0.83	-0.01	0.48	0.16
AIC	-2.88	-3.00	-4.26	-2.70	-1.38	-6.67
Average RMSE	0.46	0.54	0.70	0.43	0.38	0.64

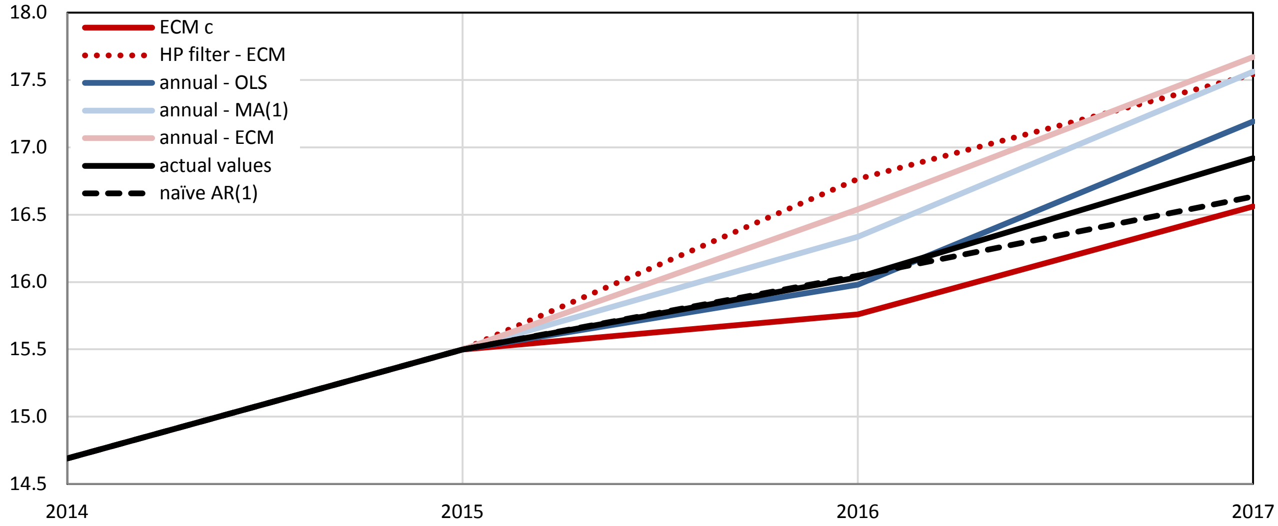
*Note: standard errors in parentheses, * p < 0.10; ** p < 0.05; *** p < 0.01. Two models based on quarterly data (Q) in the right part of the table correspond to the respective models from Table 6, only RMSE differs, due to calculation on yearly basis for comparability with RMSEs of other annual models.*

Source: Ministry of Finance calculations.

Annual time series are of course significantly shorter, which is a disadvantage of models based on annual data, especially in the case of ECM. Table 7 shows a robust dependence on nominal GDP. Regarding the forecast error, the naïve AR(1) process is surprisingly best, while the autoregressive term is statistically insignificant. Therefore, the proposed model would mean, in essence, only the assumption of a constant increase in contributions from self-employed persons of 5% per year. Compared to the results for quarterly data, however, the quarterly ECM is best, even in forecast accuracy. As the resulting model, we choose the ECM model for the non-smoothed quarterly series (red in Figures 6 and 7).

Figure 7: Comparison of Actual and Forecasted Revenues from Self-employed Payers – Annual Data

(Billions of CZK)



Source: Annual Reports of health insurance companies, Ministry of Finance calculations.

5 Persons without Taxable Income

According to Section 3b of the Act No. 592/1992 Coll., health insurance contributions for persons without taxable income are determined as a percentage of the minimum wage. The basic model for the forecast of health insurance contributions collected from persons without taxable income therefore has the following form:

$$\Delta\log(HIC\_PWTI_t) = \alpha_0 + \alpha_1 \Delta\log(MIN\_WAGE_t) + \varepsilon_t \quad (9)$$

HIC_PWTI_t is the health insurance contribution collected from persons without taxable income and MIN_WAGE_t represents the minimum wage (on which the amount of the health insurance contribution depends predominantly). Such a model additionally contains autocorrelation corresponding to the first order moving average process. However, the inclusion of moving average term does not increase the model's forecast accuracy. However, due to the nature of the time series, it is possible to consider a model without logarithms. Logarithmic transformation in this case paradoxically increases variability. Table 8 summarizes results of these specifications, including comparison with the reference AR(1) process.

Table 8: Regression Results for Revenues from Persons without Taxable Income

Contributions from persons without taxable income	OLS	MA(1)	naïve AR(1)	MA(1) no log
Constant	0.04	0.04	0.06	9.78
	(0.06)	(0.03)	(0.05)	(6.88)
Minimum wage	0.26	0.30		0.03
	(1.50)	(1.73)		(0.07)
AR(1)			-0.42 ***	
			(0.11)	
MA(1)		-0.59 ***		-0.80 ***
		(0.10)		(0.15)
Number of observations	73	73	72	73
Adj. R-squared	-0.02	0.23	0.17	0.35
AIC	1.21	0.96	1.03	12.87
Average RMSE	0.16	0.19	0.16	0.10

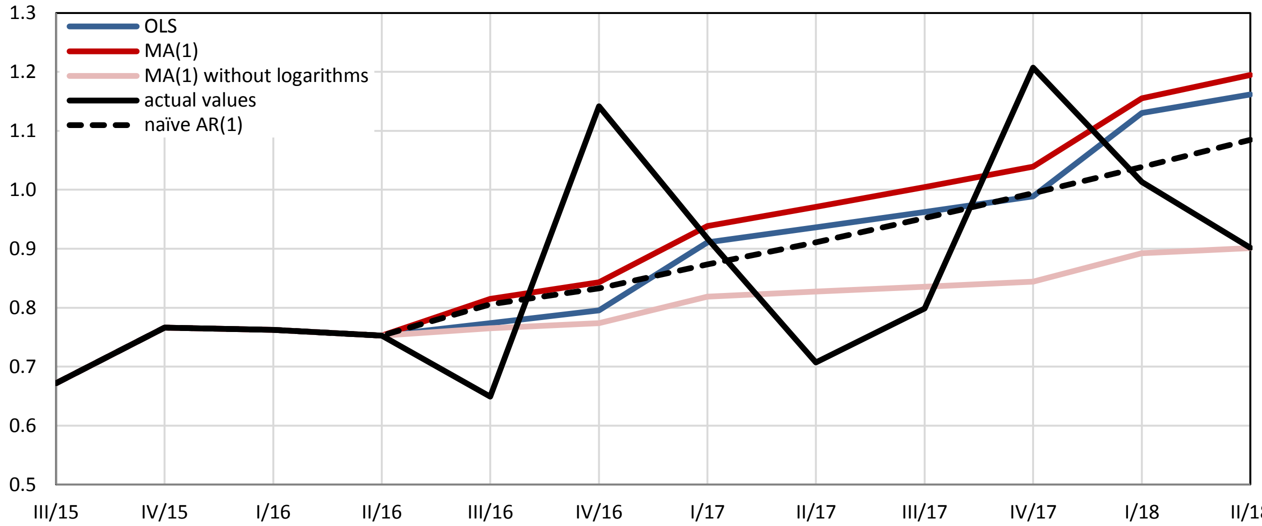
*Note: standard errors in parentheses, * p < 0.10; ** p < 0.05; *** p < 0.01. The last column represents the model without logarithmic transformation.*
Source: Ministry of Finance calculations.

The minimum wage does not explain anything about the development of health insurance contributions for persons without taxable income. However, the time series is autocorrelated, corresponding (at least in the case without logarithms) to the MA(1) process. This model also has the best forecast accuracy. The resulting model is therefore the MA(1) process without logarithmic transformation:

$$\Delta(HIC\_PWTI_t) = \alpha_0 + \alpha_1 \Delta(MIN\_WAGE_t) + \varepsilon_t + \theta\varepsilon_{t-1} \quad (10)$$

Figure 8: Comparison of Actual and Forecasted Revenues from Persons without Taxable Income

(Billions of CZK)



Source: Annual Reports of health insurance companies, Ministry of Finance calculations.

Conclusion

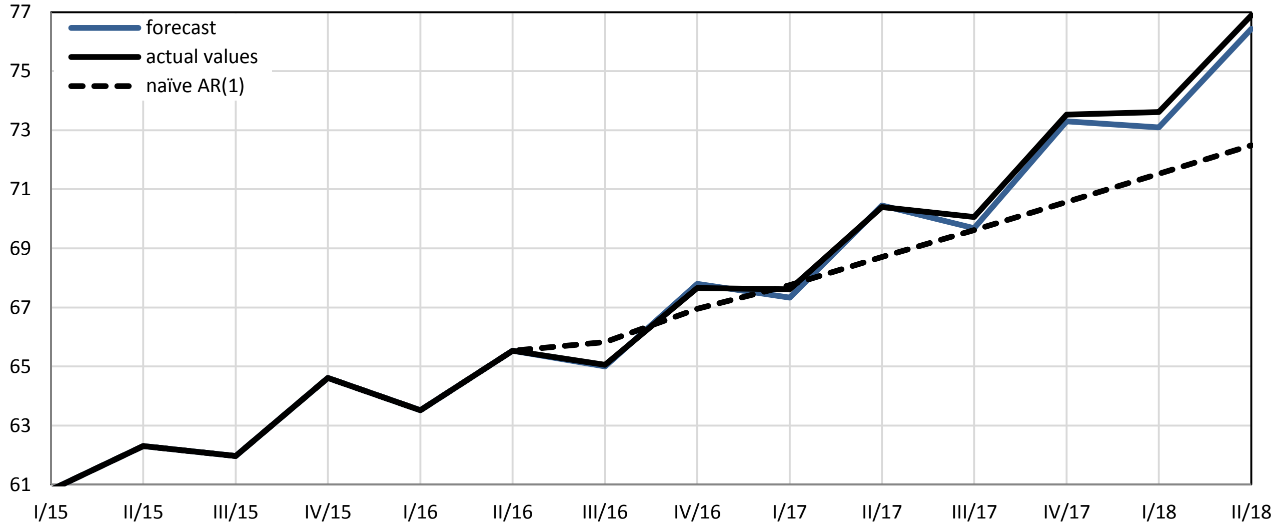
The total public health insurance revenues are almost exclusively made up of the sum of health insurance contributions from individual payer groups. Therefore, we choose the same composite approach when generating a forecast of it, that is, an aggregate of partial forecasts for individual payers. By far the largest part is made up of employee-paid health insurance contributions. Of the approx. CZK 282 billion of health insurance contributions in 2017, about 70% is from employees, 23% from state insured persons and only 6% and 1% from self-employed persons and persons without taxable income, respectively. These shares are relatively stable even in the longer term.

In summary – health insurance contribution from employees is forecasted by the model (4) FMOLS (OLS with correction of errors resulting from cointegration), state insured persons are forecasted using AR(1) process according to model (6), revenues from self-employed persons best correspond to the error correction model and finally persons without taxable income are forecasted using model (10), i.e. MA(1) without logarithmic transformation.

In Figure 9 we provide a comparison of the resulting in-sample forecast with values for the sum of naïve AR(1) forecasts. Although the resulting forecast of the Ministry of Finance shows a significantly lower RMSE, the naïve AR(1) model offers, for a certain period, relatively meaningful results. This is, however, due to lucky coincidence, as underestimation of revenues from employees is offset by a significant overestimation of revenues from state insured persons.

Figure 9: Comparison of Actual and Forecasted Total Public Health Insurance Revenues

(Billions of CZK)



Source: Annual Reports of health insurance companies, Ministry of Finance calculations.

References

Act No. 551/1991 Coll., on General Health Insurance Company of the Czech Republic, as Amended.

Act No. 280/1992 Coll., on Ministerial, Sectoral, Business and Other Health Insurance Companies, as Amended.

Act No. 592/1992 Coll., on Public Health Insurance Contribution, as Amended.

Act No. 48/1997 Coll., on Public Health Insurance, as Amended.

Act No. 23/2017 Coll., on Fiscal Responsibility Rules.

CZSO (2018a): Health Accounts of the Czech Republic - 2010–2016. Prague, Czech Statistical Office, 7 June 2018 [accessed on 19 June 2018], <https://www.czso.cz/csu/czso/vysledky-zdravotnickych-uctu-cr>.

CZSO (2018b): Gross domestic product - quarterly national accounts. Prague, Czech Statistical Office, 31 August 2018 [accessed on 10 September 2018], <https://www.czso.cz/csu/czso/hdp_ts>.

MF CR (2018): Macroeconomic Forecast of the CR, July 2018. Ministry of Finance of the CR, 31 July 2018 [accessed on 31 July 2018], <https://www.mfcr.cz/assets/en/media/Macroeconomic-Forecast-July-2018.pdf>.

OECD, Eurostat and WHO (2017): *A System of Health Accounts 2011: Revised edition*, OECD Publishing, Paris. [cit. 19. 6. 2018], <http://dx.doi.org/10.1787/9789264270985-en>.

Ministry of Finance of the Czech Republic
Letenská 15
118 10 Prague 1

http://www.mfcr.cz

ISBN 978-80-7586-023-1