Eur. Actuar. J. (2017) 7:257–276
DOI 10.1007/s13385-017-0148-0

ORIGINAL RESEARCH PAPER

Econometric model of non-life technical provisions: the Czech insurance market case study

Radek Hendrych[1] · **Tomáš Cipra**[1]

Received: 15 June 2016 / Revised: 24 November 2016 / Accepted: 4 February 2017 /
Published online: 3 March 2017
© EAJ Association 2017

Abstract The paper introduces and discusses a complex econometric model of non-life technical provisions based on the Czech non-life insurance market data. Selected economic-actuarial relations among given insurance variables are described by means of the dynamic linear system of simultaneous equations used in econometrics. In particular, the provision for outstanding claims, the provision for unearned premium, the other (marginal) technical provisions, the acquisition and administrative expenses, the benefit expenses, and their mutual interactions are studied in detail. The suggested simultaneous equations model is estimated, statistically verified, and interpreted with special regard to the actuarial point of view. The proposed modelling scheme can be further employed for prognosing the considered non-life technical provisions. Particularly, such forecasts can be taken into account by non-life insurance companies in their internal calculations (e.g. for financial planning purposes, for testing the sufficiency of non-life technical provisions, or for liability adequacy tests LAT) or by an insurance regulator or supervisory authority (e.g. for performing stress tests). Alternatively, this approach might motivate development of internal models applicable in the Solvency II framework. Both deterministic and randomly generated scenarios are analysed which can deliver relevant outputs for formulating crucial recommendations and conclusions.

Keywords Econometric system of simultaneous equations · Non-life insurance · Scenario analysis · Solvency II · Technical provisions

---

✉ Radek Hendrych
hendrych@karlin.mff.cuni.cz

Tomáš Cipra
cipra@karlin.mff.cuni.cz

[1] Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Prague 8, Czech Republic

1 Introduction

Technical provisions are undoubtedly key insurance variables. They represent the amount of money maintained by an insurance company needed to meet all its future liabilities towards the clients (under a certain measurement of present liabilities). The technical provisions must be sufficient to cover at any moment all these anticipated commitments. It should be ensured by various regulatory principles introduced, e.g. by the Solvency II regulatory system. The sufficiency is continuously monitored by the regulators and other supervisory authorities (e.g. in the Czech case by the Czech National Bank).

Generally, one distinguishes between the life and the non-life technical provisions (according to the underlying insurance contracts). All the provisions are regularly recalculated, tested, and reported in the annual (quarterly, monthly) balance sheets on the liability side. Note that there exist several exactly specified categories of the technical provisions given by the national legal framework.

The present paper introduces and discusses a complex econometric model of the most relevant non-life technical provisions based on the Czech non-life insurance market data. In particular, the dynamic linear system of simultaneous equations is employed in order to describe different interactions among selected economic-actuarial insurance variables. Namely, the provision for outstanding claims, the provision for unearned premium, and the other (marginal) technical provisions are studied in greater detail. After statistical verification and interpretation, the considered modelling scheme can be further used for prognosing the considered non-life technical provisions. These forecasts can be taken into account by non-life insurance companies in their internal calculations (e.g. for financial planning purposes, for testing the sufficiency of non-life technical provisions, or for setting the prudency level) and by national insurance regulators or other supervisory authorities (e.g. to perform stress tests or to monitor the declared prudency level). Alternatively, they might be useful for formulating internal models applicable in the Solvency II framework.

Complex econometric models, which investigated behaviour of cash flows or technical provisions in the life insurance, were discussed in various academically or practically oriented works (Baranoff, Papadopoulos, and Sager [1], Cipra [2], Feilmeier and Junker [7], Hendrych [9], Hendrych and Cipra [10]). However, to the best of our knowledge there has not been published yet any complex econometric model examining non-life technical provisions (or even based on the Czech non-life insurance market data). On the contrary, selected non-life technical provisions and related aspects have been analysed in the literature from the statistical or actuarial points of view (Dahms [5], Hürlimann [11], and many others).

The paper is organized as follows. Section 2 introduces and clarifies the proposed dynamic econometric model of non-life technical provisions for the Czech insurance market. Section 3 estimates, verifies, and interprets the model and its parameters in detail. Short-run, cumulative, and long-run effects of exogenous variables are studied and discussed. Sections 4 and 5 analyse various scenarios of future development of the particular non-life technical provisions. Simulations are based on either prescribed (deterministic) values of input (strictly exogenous) variables or their stochastically generated counterparts. In both cases, the residual bootstrap method is applied. Such prognoses might be further employed, e.g. in order to monitor the adequacy of provisions or to perform stress tests. Finally, Sect. 6 contains conclusions.

2 Model of non-life technical provisions for the Czech insurance market

As was mentioned above, we shall concentrate on the following three key categories of the non-life technical provisions: (a) the provision for outstanding claims, (b) the provision for unearned premium, and (c) the other (marginal) provisions (i.e. the sum of all other marginal non-life technical provisions representing only a minority of the total volume of all non-life provisions). The provision for outstanding claims is an estimated value of (future) compensations for policyholders and policy beneficiaries. More specifically, it involves the provision for IBNR (Incurred But Not Reported) claims and the provision for RBNS (Reported But Not Settled) claims. The provision for unearned premium corresponds to such a part of the written premium, which relates to future accounting periods. The other (but separately marginal) provisions involve, for instance, the balance provision or the provision for bonuses and sales.

For simplicity, let us consider only relationships arising from the quarterly published summary balance sheets (the liability side) of all the Czech non-life insurance companies so that the seasonality must be taken into account. Obviously one could investigate interactions among these accounting data through the econometric modelling concepts based on the actuarial theory (see e.g. [3]), and extend the introduced dataset by including other insurance or economic variables.

In particular, we assume the following non-life insurance variables: CS_t—the claims expenses in time t (in thousands of CZK), EAC_t—the acquisition expenses in time t (in thousands of CZK), EAD_t—the administrative expenses in time t (in thousands of CZK), EB_t—the existing business in time t (i.e. the number of existing non-life insurance contracts, in pieces), NRC_t—the number of reported claims in time t (in pieces), TPC_t—the technical provision for outstanding claims in time t (in thousands of CZK), TPO_t—the other non-life technical provisions in time t (in thousands of CZK), TPP_t—the technical provision for unearned premium in time t (in thousands of CZK), TPT_t—the total reported non-life technical provisions in time t (in thousands of CZK) defined as $TPT_t = TPC_t + TPO_t + TPP_t$, WP_t—the written premium in time t (in thousands of CZK), $t = 1, \ldots, 28$ ($t = 1$ refers to the Q4 2008 and $t = 28$ to the Q3 2015). The notation Q1, …, Q4 refers to the corresponding quarters of a year. The quarterly based dataset was obtained from the quarterly reported summary balance sheets (namely on the liability side) published by the Czech National Bank (ČNB) on the regular basis.[1] The dataset was available only for this period (i.e. we have not omitted any available information).

---

[1] http://www.cnb.cz/cnb/STAT.ARADY_PKG.STROM_DRILL?p_strid=BC&p_lang=CS, retrieved March 2, 2016.

We can proceed to the formulation of the dynamic linear econometric system of simultaneous equations, which describe relationships among the particular non-life insurance market variables listed above. The considered modelling scheme simultaneously explains the casual relations among more than one dependent variable. Therefore, it enables to model the analysed phenomenon in greater complexity. To be more precise, it reflects mutual interactions among the studied insurance variables through the following modelling structure (by assuming non-trivially correlated residuals).

We have considered the following system of simultaneous equations model ($t = 2, \ldots, T$):

$$\begin{aligned}
TPP_t &= \beta_{11} + \beta_{21} 1_{[Q2]} + \beta_{31} 1_{[Q3]} + \beta_{41} 1_{[Q4]} + \phi_{11} TPP_{t-1} + \beta_{61} \Delta WP_t^* + \varepsilon_t^{TPP}, \\
TPC_t &= \beta_{12} + \beta_{22} 1_{[Q2]} + \beta_{32} 1_{[Q3]} + \beta_{42} 1_{[Q4]} + \phi_{22} TPC_{t-1} + \phi_{62} CS_{t-1}^* + \beta_{82} EB_t + \varepsilon_t^{TPC}, \\
TPO_t &= \beta_{13} + \phi_{33} TPO_{t-1} + \gamma_{53} \Delta EAD_t^* + \phi_{43} EAC_{t-1}^* + \beta_{73} NRC_t^* + \varepsilon_t^{TPO}, \\
EAC_t^* &= \beta_{14} + \phi_{14} TPP_{t-1} + \beta_{94} EB_{t-1} + \beta_{64} WP_{t-1}^* + \beta_{74} NRC_t^* + \varepsilon_t^{EAC}, \\
EAD_t^* &= \beta_{15} + \beta_{55} \Delta WP_t^* + \phi_{55} EAD_{t-1}^* + \beta_{85} EB_t + \phi_{35} TPO_{t-1} + \varepsilon_t^{EAD}, \\
CS_t^* &= \beta_{16} + \phi_{66} CS_{t-1}^* + \beta_{76} NRC_t^* + \beta_{56} WP_t^* + \beta_{86} EB_t + \varepsilon_t^{CS}, \\
TPT_t &= TPC_t + TPO_t + TPP_t,
\end{aligned} \tag{1}$$

where $1_{[\cdot]}$ denotes the binary indicator of the event $\cdot$, Δ stands for the first difference operator, and the superscript * indicates that the variable has been seasonally adjusted (by using a simple routine multiplicative seasonal factor method [4]). Moreover, β's, γ's, and ϕ's with various indices represent the unknown parameters of the model and ε's denote the stochastic error terms.

The considered dynamic econometric system (1) includes six stochastic equations (i.e. the equations with the stochastic residual terms) and one deterministic equation (i.e. the definition equation for the total provisions *TPT*). In the suggested model, the intercept, the seasonal dummies, and the variables *EB*, NRC^*, WP^* (and thus also the lagged *EB* and WP^*) are assumed to be strictly exogenous (i.e. uncorrelated with residual components at all times, since these variables enter into the system from outside). Such a particular choice of exogenous variables seems to be pragmatic with regard to the apparent external character of these variables. Furthermore, the lagged endogenous variables CS^*, EAC^*, EAD^*, *TPC*, *TPO*, and *TPP* are supposed to be predetermined (i.e. uncorrelated with current and future residual disturbances since they are fully determined by the system (1) in time t - 1). To sum up, the model (1) includes six endogenous, nine strictly exogenous, and six predetermined variables. Note that each equation in (1) satisfies the necessary order condition of identification [4, 7], i.e. the number of the variables on the right-hand side of each equation is less or equal to the number of exogenous variables in the system. In particular, all the equations in (1) are overidentified (i.e. the number of exogenous variables excluded from the given equation is greater than the number of endogenous variables included on the right-hand side of the equation). This is consistent with the common econometric practice that the most of simultaneous equations are overidentified.

To be more precise, the model (1) after ignoring the last definition equation follows the structural form of the dynamic system of linear simultaneous equations (see e.g. [8]):

$$\mathbf{y}_t^{\mathrm{T}}\boldsymbol{\Gamma} + \mathbf{y}_{t-1}^{\mathrm{T}}\boldsymbol{\Phi}_1 + \mathbf{x}_t^{\mathrm{T}}\mathbf{B} + \boldsymbol{\varepsilon}_t^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}, \tag{2}$$

where

$$\begin{aligned} \mathbf{y}_t &= (TPP_t, TPC_t, TPO_t, EAC_t^*, EAD_t^*, CS_t^*)^{\mathrm{T}}, \\ \mathbf{x}_t &= (1, 1_{[Q2]}, 1_{[Q3]}, 1_{[Q4]}, WP_t^*, WP_{t-1}^*, NRC_t^*, EB_t, EB_{t-1})^{\mathrm{T}}, \\ \boldsymbol{\varepsilon}_t &= (\varepsilon_t^{TPP}, \varepsilon_t^{TPC}, \varepsilon_t^{TPO}, \varepsilon_t^{EAC}, \varepsilon_t^{EAD}, \varepsilon_t^{CS})^{\mathrm{T}}, \end{aligned} \tag{3}$$

where $\mathbf{y}_t$ denotes the (6×1) vector of endogenous variables, $\mathbf{x}_t$ is the (9×1) vector of strictly exogenous variables, and $\boldsymbol{\varepsilon}_t$ stands for the (6×1) stochastic vector of structural error terms (everything for all t). Point out that the indices of the parameters β, γ, and ϕ in (1) refer to the corresponding elements of the parameter matrices $\mathbf{B}$ (9×6) concerning strictly exogenous variables, $\boldsymbol{\Gamma}$ (6×6) concerning endogenous variables, and $\boldsymbol{\Phi}_1$ (6×6) concerning predetermined variables with the unit time lag, respectively. Moreover, some a priori constraints must be introduced, i.e. $\beta_{61} = -\beta_{51}$, $\gamma_{53} = -\phi_{53}$, $\beta_{65} = -\beta_{55}$. Other elements of these matrices of parameters are equal to zero. Several other assumptions are usually introduced: (A1) $\mathrm{E}(\boldsymbol{\varepsilon}_t) = \mathbf{0}$ for all t, $\mathrm{var}(\boldsymbol{\varepsilon}_t) = \mathrm{E}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t^{\mathrm{T}}) = \boldsymbol{\Sigma}$ is a finite symmetric positive definite matrix constant for all t, and finally $\mathrm{cov}(\boldsymbol{\varepsilon}_s, \boldsymbol{\varepsilon}_t) = \mathrm{E}(\boldsymbol{\varepsilon}_s\boldsymbol{\varepsilon}_t^{\mathrm{T}}) = \mathbf{0}$ for all $s \neq t$; (A2) $\mathrm{E}((\mathbf{y}_{t-1}^{\mathrm{T}}, \mathbf{x}_t^{\mathrm{T}})^{\mathrm{T}}(\mathbf{y}_{t-1}^{\mathrm{T}}, \mathbf{x}_t^{\mathrm{T}})) = \mathbf{Q}$ is a finite symmetric positive definite matrix constant for all t; (A3) the matrix $\boldsymbol{\Gamma}$ containing parameters concerning (exclusively) endogenous variables is invertible with elements -1 on its diagonal. For more details, consult [8] or [12].

3 Model estimation, verification, and interpretation

To estimate the unknown parameters of the proposed simultaneous equations model (1), the three-stage least squares method (3SLS) might be applied. This full information estimation technique is a special case of the generalized method of moments GMM exploiting all information available in the considered system. It guarantees suitable properties (under general assumptions). Namely, the 3SLS estimates are consistent, asymptotically normally distributed, and asymptotically efficient [8]. Note that the 3SLS estimation procedure is easily accessible in common econometric software, e.g. EViews, R, or Stata. It can be handled comfortably and does not depend on distribution of the error terms. Table 1 presents the 3SLS estimates of the model (1) jointly with the estimated standard errors and the achieved coefficients of determination R^2. The goodness of fit results for *EAD* is different than all others. This is likely caused by market turbulences in the period 2011–2012 (see Fig. 1). Table 2 reports the 95% bootstrap studentized confidence

Table 1 The 3SLS estimates of the parameters of the model (1)

Eq. for TPP		Eq. for TPC		Eq. for TPO		Eq. for EAC*		Eq. for EAD*		Eq. for CS*	
R^2	98.93%	R^2	92.90%	R^2	98.37%	R^2	98.13%	R^2	83.80%	R^2	92.14%
Par.	Est. (Std. Err.)	Par.	Est. (Std. Err.)	Par.	Est. (Std. Err.)	Par.	Est. (Std. Err.)	Par.	Est. (Std. Err.)	Par.	Est. (Std. Err.)
β_{11}	1467464 (451313)	β_{12}	2039251 (2796570)	β_{13}	14165828 (2355754)	β_{14}	4799534 (565896)	β_{15}	447354 (517488)	β_{16}	−6429938 (2272260)
β_{21}	−693205 (201651)	β_{22}	833827 (404924)	β_{73}	−4.48788 (1.47497)	β_{64}	−0.06710 (0.01208)	β_{55}	0.11712 (0.01581)	β_{56}	0.31614 (0.05910)
β_{31}	−1831019 (194997)	β_{32}	1437856 (405270)	γ_{53}	1.32742 (0.25785)	β_{74}	1.22320 (0.35485)	β_{85}	0.08230 (0.02454)	β_{76}	15.16822 (2.02663)
β_{41}	−1134179 (199541)	β_{42}	618533 (418687)	ϕ_{33}	0.54494 (0.06365)	β_{94}	0.12381 (0.02003)	ϕ_{35}	0.04138 (0.01479)	β_{86}	−0.55251 (0.10278)
β_{61}	0.37131 (0.08208)	β_{82}	0.92553 (0.25469)	ϕ_{43}	−0.89766 (0.19209)	ϕ_{14}	0.09357 (0.00774)	ϕ_{55}	0.56685 (0.10284)	ϕ_{66}	0.70127 (0.06974)
ϕ_{11}	0.98911 (0.02169)	ϕ_{22}	0.69778 (0.07629)								
		ϕ_{62}	−0.26473 (0.14110)								

Source: authors (by EViews 8.0)

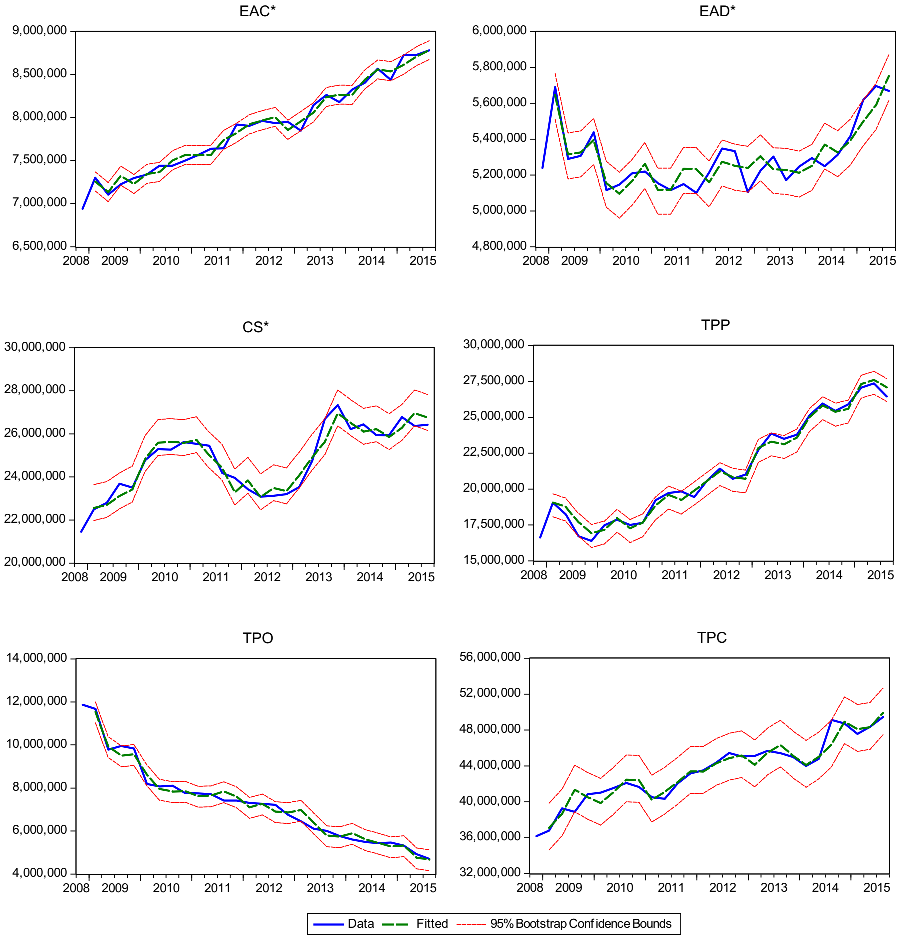

Fig. 1 The observed endogenous variables with their fitted counterparts together with the corresponding 95% bootstrap confidence bounds (source: authors by EViews 8.0)

intervals for all the parameters of the model (1). See e.g. [10] for more details. One can identify that the confidence bounds are relatively broad. According to these confidence limits, several parameters seem to be insignificant (i.e. those with the confidence intervals containing 0). Such results are in contrast with the conclusions based on the calculated z-statistics (i.e. the ratio between the estimated parameter and its estimated standard deviation), which confirm the significance of most parameters (except for selected intercepts, β_{42}, and ϕ_{62}). However, the z-statistics are evaluated by employing the asymptotic properties of 3SLS estimation, which might be inefficient for small samples.

One can see that the estimated model fits the data suitably (see Fig. 1). Figure 1 displays the original data, the corresponding fitted values, and the 95% confidence bounds computed by the bootstrap procedure as is implemented in EViews 8.0.

Table 2 The 95% bootstrap studentized confidence bounds for the parameters of the model (1)

Eq. for TPP		Eq. for TPC		Eq. for TPO		Eq. for EAC*		Eq. for EAD*		Eq. for CS*	
Par.	Lower B. Upper B.	Par.	Lower B. Upper B.	Par.	Lower B. Upper B.	Par.	Lower B. Upper B.	Par.	Lower B. Upper B.	Par.	Lower B. Upper B.
β_{11}	1385105	β_{12}	−6489212	β_{13}	−5091021	β_{14}	−509007	β_{15}	−10190360	β_{16}	−12936645
	2033686		14622077		23532155		9410131		5674406		1167070
β_{21}	−1624366	β_{22}	−1197969	β_{73}	−11.28781	β_{64}	−0.17473	β_{55}	−0.22434	β_{56}	0.08748
	517254		3594378		6.43065		0.08010		0.42858		0.55530
β_{31}	−3051514	β_{32}	−2377317	γ_{53}	−0.94974	β_{74}	−1.99711	β_{85}	−0.30366	β_{76}	6.71611
	213194		5503778		6.18829		5.05548		0.57177		27.65568
β_{41}	−2086938	β_{42}	−1495693	ϕ_{33}	0.22479	β_{94}	−0.05512	ϕ_{35}	−0.19104	β_{86}	−1.11236
	354687		3101319		1.14398		0.28079		0.32646		−0.30675
β_{61}	−0.41947	β_{82}	−0.53770	ϕ_{43}	−1.73713	ϕ_{14}	0.03164	ϕ_{55}	−1.25175	ϕ_{66}	0.34031
	1.03323		1.87716		0.55693		0.13759		2.92616		1.10542
ϕ_{11}	0.95861	ϕ_{22}	0.36497								
	0.99257		1.31956								
		ϕ_{62}	−1.23191								
			0.33298								

Source: authors (by EViews 8.0)

Moreover, the model demonstrates its statistical adequacy (refer to the assumptions listed above). Firstly, the sample correlation matrix of the estimated 3SLS residuals demonstrates several relatively high correlations (see Table 3). It points to the appropriateness of the suggested modelling framework; compare with the assumption (A1). Secondly, neither the multivariate Ljung-Box test nor the empirical autocorrelation functions of the 3SLS residuals indicate the presence of residual autocorrelations [12]. Thirdly, the joint Jarque–Bera test based on the Cholesky decomposition of the estimated covariance matrix $\mathbf{\Sigma}$ cannot reject the multivariate normality of the 3SLS residuals (the achieved p-value equals 0.108). Fourthly, the Hausman specification test comparing the two-stage and the three-stage least squares estimates delivers the statistics $W = 7.775$ with 33 degrees of freedom and p-value 0.999, i.e. the proper model specification cannot be rejected. Finally, the Sargan test for overidentifying restrictions is calculated (see Table 4). It verifies that in each equation all the exogenous variables and the residual term satisfy the orthogonality condition. A large value of the test statistics is obviously taken as evidence that there likely exist inappropriately omitted exogenous variables in the investigated equation [4, 7]. Using the calculated p-values (see Table 4), one cannot reject the correct specification of any equation in the system (1). To conclude, the suggested simultaneous equations model (1) seems to be correctly specified from the statistical point of view. Nevertheless, the results of formal statistical testing remain rather indicative due to the relatively low number of studied observations. The 5% significance level has been applied.

To clarify the variable interactions in the model (1), one conveniently uses the reduced form of the modelling scheme (2), i.e. the relation (1) multiplied from the right by the inverted matrix $\mathbf{\Gamma}^{-1}$:

$$\mathbf{y}_t^{\mathrm{T}} = -\mathbf{y}_{t-1}^{\mathrm{T}}\mathbf{\Phi}_1\mathbf{\Gamma}^{-1} - \mathbf{x}_t^{\mathrm{T}}\mathbf{B}\mathbf{\Gamma}^{-1} - \boldsymbol{\varepsilon}_t^{\mathrm{T}}\mathbf{\Gamma}^{-1},\ t = 2, \ldots, T. \tag{4}$$

This can be rewritten as follows:

$$\mathbf{y}_t^{\mathrm{T}} = \mathbf{y}_{t-1}^{\mathrm{T}}\mathbf{\Delta}_1 + \mathbf{x}_t^{\mathrm{T}}\mathbf{\Pi} + \mathbf{v}_t^{\mathrm{T}},\ t = 2, \ldots, T, \tag{5}$$

where we put:

Table 3 The estimated 3SLS residual correlation matrix

	TPP	*TPC*	*TPO*	*EAC**	*EAD**	*CS**
TPP	1.00000	0.26336	−0.39427	−0.01615	0.02415	−0.11868
TPC	0.26336	1.00000	−0.45937	0.06458	0.05888	−0.52206
TPO	−0.39427	−0.45937	1.00000	−0.09768	0.03929	0.32066
*EAC**	−0.01615	0.06458	−0.09768	1.00000	0.12879	0.17445
*EAD**	0.02415	0.05888	0.03929	0.12879	1.00000	−0.25189
*CS**	−0.11868	−0.52206	0.32066	0.17445	−0.25189	1.00000

Source: authors (by EViews 8.0)

Table 4 The Sargan statistics based on 2SLS residuals of the model (1)

Equation	Statistics	*df*	*p*-value
TPP	15.30790	8	0.05343
TPC	9.95314	8	0.26833
TPO	10.17593	9	0.33643
*EAC**	10.48992	10	0.39861
*EAD**	14.36301	9	0.10999
*CS**	11.20828	10	0.34152

Source: authors (by EViews 8.0)

$$\boldsymbol{\Delta}_1 = -\boldsymbol{\Phi}_1\boldsymbol{\Gamma}^{-1},\ \boldsymbol{\Pi} = -\mathbf{B}\boldsymbol{\Gamma}^{-1},\ \mathbf{v}_t^{\mathrm{T}} = -\boldsymbol{\varepsilon}_t^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}.$$

By using the backward substitution, one obtains:

$$\mathbf{y}_t^{\mathrm{T}} = \mathbf{y}_1^{\mathrm{T}}\boldsymbol{\Delta}_1^{t-1} + \sum_{s=0}^{t-2}\mathbf{x}_{t-s}^{\mathrm{T}}\boldsymbol{\Pi}\boldsymbol{\Delta}_1^{s} + \sum_{s=0}^{t-2}\mathbf{v}_{t-s}^{\mathrm{T}}\boldsymbol{\Delta}_1^{s},\ t = 2,\ldots,T. \tag{6}$$

This statement simply assesses how the variables on the right-hand side of (6) contribute to the change of the current values of the endogenous variables involved in **y**. In general, three types of effects are distinguished: (*a*) a short-run effect, (*b*) a cumulative effect, and (*c*) a long-run effect.

Firstly, the short-run effect of a strictly exogenous explanatory variable x_i on an endogenous variable y_j is defined by means the matrix $\boldsymbol{\Pi}$ elements. It concerns the immediate change of the studied endogenous variable as:

$$\frac{\partial y_{jt}}{\partial x_{it}} = \boldsymbol{\Pi}_{ij}. \tag{7}$$

Secondly, the cumulative effect of a strictly exogenous explanatory variable x_i on an endogenous variable y_j can be expressed as the following finite sum:

$$\sum_{s=0}^{S}\frac{\partial y_{jt}}{\partial x_{i,t-s}} = \sum_{s=0}^{S}\left(\boldsymbol{\Pi}\boldsymbol{\Delta}_1^{s}\right)_{ij},\ S = 0, 1, \ldots \tag{8}$$

Thirdly, the long-run effect of a strictly exogenous explanatory variable x_i on an endogenous variable y_j is defined in terms of the following infinite sum with respect to (6) by assuming $t \to \infty$:

$$\sum_{s=0}^{\infty}\frac{\partial y_{jt}}{\partial x_{i,t-s}} = \sum_{s=0}^{\infty}\left(\boldsymbol{\Pi}\boldsymbol{\Delta}_1^{s}\right)_{ij}. \tag{9}$$

This approach corresponds, in fact, to the impulse response analysis when applying unit impulses to the (selected) strictly exogenous variables. Moreover, if all the eigenvalues of $\boldsymbol{\Delta}_1$ lie inside the unit circle, the sum (9) can be simplified as:

$$\sum_{s=0}^{\infty} \mathbf{\Pi}\mathbf{\Delta}_1^s = \mathbf{\Pi}\left[\mathbf{I} + \mathbf{\Delta}_1 + \mathbf{\Delta}_1^2 + \ldots\right] = \mathbf{\Pi}[\mathbf{I} - \mathbf{\Delta}_1]^{-1}. \quad (10)$$

In the case of the proposed econometric system (1) and its 3SLS estimates, the absolute values of eigenvalues of the matrix $\mathbf{\Delta}_1$ are consecutively 0.989, 0.701, 0.698, 0.603, 0.603, and 0. All these eigenvalues obviously lie inside the unit circle. Therefore, the simultaneous equations model (1) can be regarded as stable from this viewpoint. It particularly means that all the discussed effects can be calculated (see results in Tables 5, 6). In addition, they provide various relevant interpretations.

For instance, we shall investigate the effects on the technical provision for outstanding claims (*TPC*) from this viewpoint since the technical provision for outstanding claims represents the major part of the total non-life technical provisions *TPT*. From Table 5, if the actual value of existing business (*EB*) is increased by 1, ceteris paribus, the expected change of *TPC* will be +0.926, i.e. a new non-life insurance contract brings extra 926 CZK into the technical provision for outstanding claims forthwith. Other strictly exogenous variables have no short-run (immediate) impact on the current value of *TPC*. On the other hand, the analysis of the long-run effects is more complex (see Table 6). If the seasonally adjusted written premium (WP^*) is increased by 1 at each time in the past, i.e. by 1000 CZK, ceteris paribus, the anticipated long-run change of *TPC* will be −0.927, i.e. the current *TPC* will be totally reduced by 927 CZK, which is relatively negligible. It follows from the interaction between $WP^*$and CS^*: the increase in WP^* produces the proportional increase of CS^*, and the increase of CS^* results in the decrease of *TPC* (seen from the long-time horizon perspective). For more details, consult the model (1) and the corresponding parameter estimates reported in Tables 1 and 2 (e.g. $\phi_{62} = -0.265$). Furthermore, if the seasonally adjusted number of reported claims (NRC^*) increases by 1 at each time in the past, ceteris paribus, the current *TPC* will be expectedly reduced by 44.475, i.e. by 44475 CZK. This amount can be regarded as the cumulative sum of average paid claims from the provision for

Table 5 Short-run effects, i.e. the 3SLS estimate of $\mathbf{\Pi}$

	TPP	*TPC*	*TPO*	EAC^*	EAD^*	CS^*
1	1467463.487	2039250.801	14759653.808	4799533.949	447353.931	−6429937.842
$1_{[Q2]}$	−693205.436	833826.810	0.000	0.000	0.000	0.000
$1_{[Q3]}$	−1831019.273	1437855.861	0.000	0.000	0.000	0.000
$1_{[Q4]}$	−1134178.795	618532.978	0.000	0.000	0.000	0.000
WP^*	0.371	0.000	0.155	0.000	0.117	0.316
WP^*_{-1}	−0.371	0.000	−0.155	−0.067	−0.117	0.000
NRC^*	0.000	0.000	−4.488	1.223	0.000	15.168
EB	0.000	0.926	0.109	0.000	0.082	−0.553
EB_{-1}	0.000	0.000	0.000	0.124	0.000	0.000

Source: authors (by EViews 8.0)

Table 6 Long-run effects, i.e. the 3SLS estimate of $\mathbf{\Pi}[\mathbf{I}\text{-}\mathbf{\Delta}_1]^{-1}$

	TPP	*TPC*	*TPO*	EAC^*	EAD^*	CS^*
1	134787469.666	25600989.618	−3216568.281	17411482.135	725524.596	−21524057.191
$1_{[Q2]}$	−63671367.301	2758977.347	11752199.897	−5957675.349	1122605.953	0.000
$1_{[Q3]}$	−168180303.536	4757596.781	31042030.817	−15736487.075	2965229.395	0.000
$1_{[Q4]}$	−104175055.318	2046610.222	19228204.549	−9747570.774	1836736.703	0.000
WP^*	34.105	−0.927	−6.295	3.191	−0.331	1.058
WP^*_{-1}	−34.105	0.000	6.427	−3.258	0.344	0.000
NRC^*	0.000	−44.475	−12.275	1.223	−1.173	50.775
EB	0.000	4.682	0.000	0.000	0.190	−1.850
EB_{-1}	0.000	0.000	−0.244	0.124	−0.023	0.000

Source: authors (by EViews 8.0)

outstanding claims determined by the newly reported claims. Moreover, if the value of existing business (*EB*) is increased by 1 at each time in the past, ceteris paribus, the expected long-run change of *TPC* will be +4.682, i.e. these new non-life insurance contracts bring extra 4682 CZK into the technical provision for outstanding claims from the long-run point of view. Analogous interpretations can be derived also for other endogenous variables.

Finally, we shall study the cumulative effects of all non-deterministic strictly exogenous variables on the total reported non-life insurance technical provisions (*TPT*). These effects are simply calculated as the sum of the cumulative effects on *TPP*, *TPC*, and *TPO*. This analysis is performed by using the formula (8). Fig. 2 displays the cumulative effects (together with the corresponding 95% bootstrap Hall confidence limits) of all non-deterministic strictly exogenous variables on *TPT*

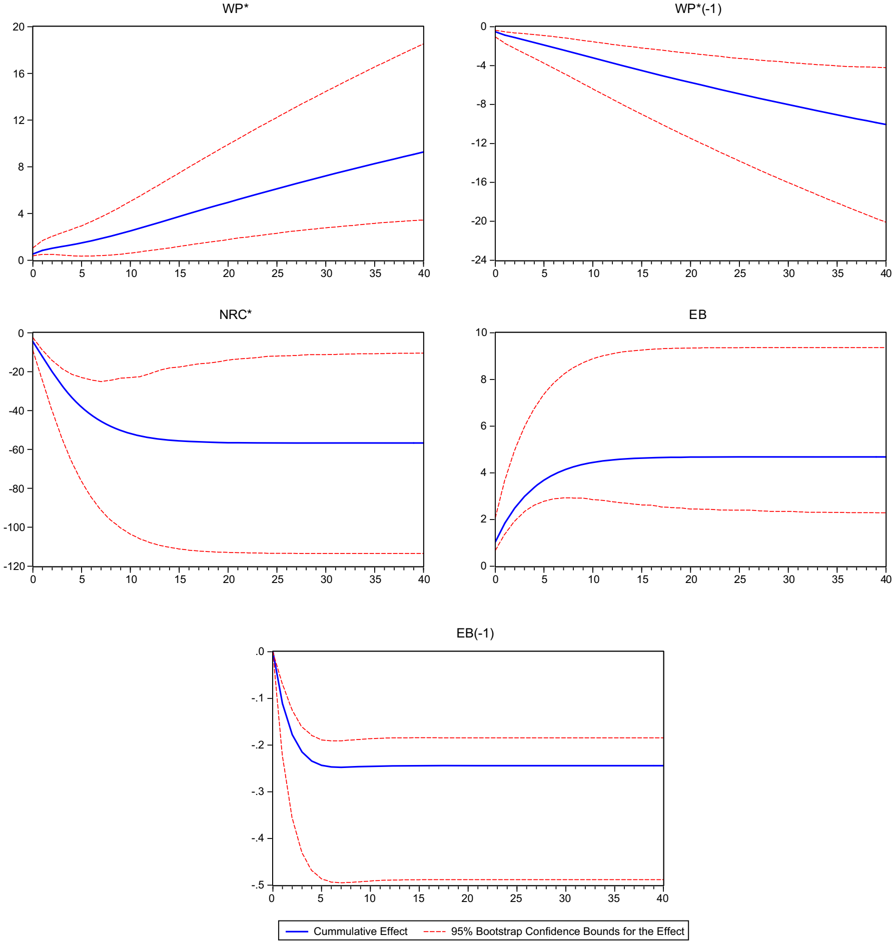

Fig. 2 The cumulative effects of all the strictly exogenous variables on the total non-life technical provisions *TPT* (source: authors by EViews 8.0)

consecutively for $S = 0,\ldots, 40$. Note that analogous graphs are usually presented when analysing impulse response functions, e.g., for a vector autoregression. It is evident that the cumulative effects gradually converge to the long-run effects, which is in accordance with the eigenvalues of $\boldsymbol{\Delta}_1$ (i.e. with the model stability). If we combine the impacts of WP^* and EB together with their lagged versions [which is necessary because they affect TPT simultaneously due to (1)], we will conclude that: (a) WP^* has a relatively negligible (negative) cumulative impact on TPT, (b) the permanent unit increase in EB at each past instant has a positive impact on TPT, (c) the most substantial negative cumulative impact on TPT is observed in the case of NRC^*. This could be anticipated since the number of reported claims NRC^* is supposed to affect the non-life technical provisions considerably (especially, the technical provision for outstanding claims as can be recognized from Table 6).

4 Scenario analysis: prescribed strictly exogenous variables

From the practical point of view, prognosing the considered non-life technical provisions might be truly useful. For instance, one could employ randomly generated scenarios for stress testing the adequacy of the particular non-life technical provisions [by applying the modelling scheme (5)]. Consequently, this approach might motivate development or verification of internal models for non-life insurance companies, which may be introduced in the Solvency II framework (to verify in a parallel way the best estimates of technical provisions, to have an efficient instrument for stress testing or scenario generation, and other possibilities). Moreover, it could be helpful for any supervisory authority to set or to monitor the prudency level effectiveness. Undoubtedly, each particular application of such an econometric model should be thoroughly examined in order to respect the regulatory framework as is described in [6].

To illustrate the key idea of the stress testing discussed above, three different stress scenarios for the strictly non-deterministic exogenous variables EB, NRC^*, and WP^* are considered. All scenarios are rather pessimistic since one usually tests the sufficiency (or the prudency level effectiveness) of the technical provisions under (extremely) unfavourable conditions. The first scenario is formulated as follows: the number of existing non-life insurance contracts EB and the written premium (seasonally adjusted) WP^* remain as in Q3 2015 for the whole prediction horizon, the number of reported claims (seasonally adjusted) NRC^* increases by 3% each quarter, $t = 29, \ldots, 37$, i.e. from Q4 2015 to Q4 2017. The second scenario is described as follows: the number of existing non-life insurance contracts EB decreases by 3% each quarter, the number of reported claims (seasonally adjusted) NRC^* increases by 3% each quarter, and the written premium (seasonally adjusted) WP^* decreases by 5% each quarter, $t = 29, \ldots, 37$, i.e. again from Q4 2015 to Q4 2017. Note that the seasonally adjusted written premium decreases faster than the existing business. The third underlying scenario follows analogous expectations as the previous one introducing only minor changes: the number of existing non-life insurance contracts EB decreases by 5% each quarter, the number of reported claims (seasonally adjusted) NRC^* increases by 3% each quarter, and the written premium (seasonally adjusted) WP^* decreases by 3% each quarter, $t = 29, \ldots, 37$, i.e. from Q4 2015 to Q4 2017. Here on the contrary, the seasonally adjusted written premium decreases more slowly than the existing business portfolio.

Accepting these three stress scenarios, we can further employ the estimated modelling scheme (5) (see the outputs displayed in Tables 1, 2). In particular, we have calculated 10000 realizations (forecasts) of all the endogenous variables for each stress scenario and the whole prediction horizon by using the prescribed development of strictly exogenous variables EB, NRC^*, and WP^* and 10000 randomly generated vector error terms $\boldsymbol{\varepsilon}_t$. We have applied the standard residual bootstrap method as in [10]. It particularly means that the multivariate distribution of the disturbances $\boldsymbol{\varepsilon}_t$ is determined by the (centred) empirical residuals computed during the realized 3SLS estimation. All computations were performed in EViews version 8.0 by authors' own procedures. It is noteworthy that this analysis could be extended in order to reflect financial risks (e.g. discounting factors) to be market-consistent as is stated in [6]. On the other hand, the prediction horizon is relatively short; therefore, we have neglected this requirement (we shall study only the future values of analysed variables in Sects. 4 and 5). For longer horizons, one can incorporate discounting factors or interest rates explicitly into the stochastic generator as a next level of the model.

The simulation study results are summarized in Figs. 3 and 4. At first sight, one can see that all the generated scenarios have an impact on all the presented non-life technical provisions. Nevertheless, the results correspond to our rational anticipations with regard to the considered scenarios. Furthermore, one can identify that the

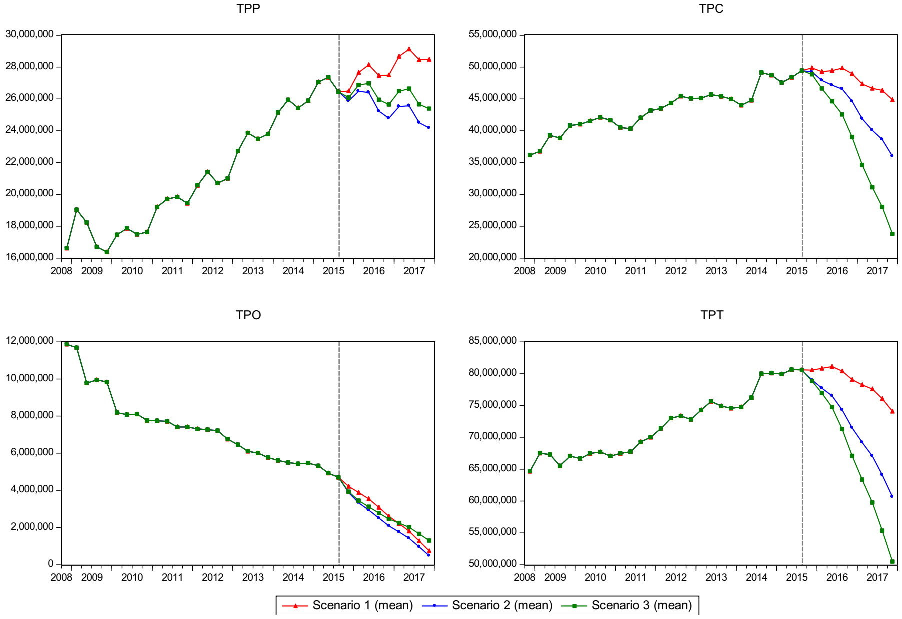

Fig. 3 The forecasts of particular non-life technical provisions in stress testing: scenario analysis with prescribed strictly exogenous variables (source: authors by EViews 8.0)

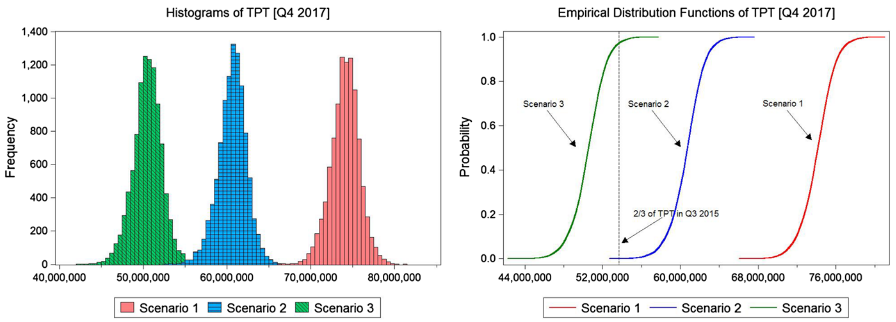

Fig. 4 The forecast of the total non-life technical provisions *TPT* for Q4 2017 in stress testing: scenario analysis with prescribed strictly exogenous variables (source: authors by EViews 8.0)

differences among all the stress scenarios are truly significant. To be more specific, compare the results of prognosing the total non-life technical provisions (*TPT*) in Fig. 3. Here, the (mean) forecasted provision *TPT* is evidently the lowest under the third underlying stress scenario (comparing with the first and second ones). The third stress scenario leads to the most unfavourable outcomes from this perspective. However, it perfectly correlates with the examined cumulative effect of the strictly exogenous variables on the particular non-life technical provisions. From the analysis of the cumulative effect, which is delivered in Sect. 3, one presumes that the number of existing non-life insurance contracts (*EB*) and the number of reported claims (NRC^*) considerably affect the non-life technical provisions (consult Fig. 2).

From the actuarial viewpoint, the generated scenario projections provide several useful interpretations (see Fig. 4). For instance, we can observe that the empirical probability that the total non-life technical provision *TPT* in Q4 2017 will be greater than the two-thirds of the Q3 2015 level is 100% for the first stress scenario, 99.95% for the second stress scenario, and only 2.54% for the third stress scenario, respectively. The substantial differences among the proposed scenarios are indeed apparent. Such outputs could be further employed, e.g. by the regulator for testing the sufficiency of the total provisions or for calibrating the prudency level. Table 7 reports estimated Value at Risk (*VaR*) and Expected Shortfall (*ES*) measures (at the formal confidence level $\alpha = 0.5\%$, which corresponds to "1 in 200 event" in the Solvency II lingo). It should be highlighted that we introduce the estimated VaR and ES values for the left tail of technical provisions distribution (for a particular

Table 7 Value at Risk and Expected Shortfall measures (the confidence level $\alpha = 0.005$) for the predicted total non-life technical provisions *TPT* in Q4 2017: scenario analysis with prescribed strictly exogenous variables

TPT Q4 2017	Scenario 1	Scenario 2	Scenario 3
VaR	69357840	55585509	45578637
ES	68519329	54886307	44894427

Source: authors (by EViews 8.0)

prediction horizon) since we study the sufficiency of future technical provisions (we appraise the less favourable scenarios). This also implies that the estimated ES is always smaller than VaR. The Value at Risk and Expected Shortfall measures are estimated non-parametrically as follows (i.e. without any explicit assumption on distribution):

$$VaR(\alpha) = \hat{F}^{-1}(\alpha) \text{ and } ES(\alpha) = \frac{1}{\alpha}\int_0^{\alpha} \hat{F}^{-1}(u)\mathrm{d}u,\ \alpha \in (0,1), \tag{11}$$

where $\hat{F}^{-1}(\alpha)$ denotes the estimated (continuous) quantile function of a particular technical provision (for a particular prediction horizon).

From these outputs, it is evident that Scenario 3 is the less favourable for the Czech non-life insurance market from the considered scenario portfolio.

5 Scenario analysis: randomly generated strictly exogenous variables

In Sect. 4, we have analysed behaviour of the non-life technical provisions by simulations based on three different deterministic (prescribed) scenarios for the strictly exogenous variables EB, NRC^*, and WP^*. This approach for testing technical provisions is undoubtedly very frequent in practice. On the contrary, it seems also useful to investigate the considered non-life technical provisions by employing stochastically generated strictly exogenous variables. Outputs of such an analysis may be employed analogously as before, i.e. for verifying the adequacy of technical provisions, for financial planning purposes, etc.

In particular, we assume that the strictly exogenous variables (WP^*, NRC^*, and EB) are generated by the multivariate random walk with drift for $t = 29, \ldots, 37$, i.e. again from Q4 2015 to Q4 2017. Namely, the variables respect the following modelling scheme:

$$\begin{pmatrix} WP_t^* \\ NRC_t^* \\ EB_t \end{pmatrix} = \boldsymbol{\mu} + \begin{pmatrix} WP_{t-1}^* \\ NRC_{t-1}^* \\ EB_{t-1} \end{pmatrix} + \xi_t, \tag{12}$$

where $\boldsymbol{\mu}$ is the (3×1) vector of constants (i.e. the drift) and ξ_t denotes the multivariate (strict) white noise with zero mean and the (3×3) covariance matrix $\boldsymbol{\Omega}$. Note that the vector $\boldsymbol{\mu}$ is estimated by the common OLS method based on the given sample from $t = 2$ to $t = 27$. All strictly exogenous variables are thus simulated by using the model (12), where $\boldsymbol{\mu}$ is substituted by its OLS estimate and where the error disturbances ξ_t are determined by the residual bootstrap method discussed in Sect. 4 (see also [10]). To be more precise, for $t = 29, \ldots, 37$ the values $\{\xi_t\}$ are randomly drawn (without replacement) from the centred OLS residuals of the scheme (12) estimated for $t = 2$ to $t = 27$ (see also above). In particular, it means that ξ_t is distributed in a non-standard way.

After having generated one realization of the strictly exogenous variables by employing the above mentioned random walk process, we have calculated one realization of all the endogenous variables given by (5) for the whole prediction horizon by using the values of the simulated strictly exogenous variables EB, NRC^*, and WP^* and the vector error terms $\boldsymbol{\varepsilon}_t$ randomly generated by the residual bootstrap scheme. In particular, for $t = 29, \ldots, 37$ the values $\{\boldsymbol{\varepsilon}_t\}$ are randomly drawn (without replacement) from the centred 3SLS residuals calculated in Sect. 3 (see also Sect. 4). We have repeated this process ten thousand times. Apparently, this simulation scheme has combined two independent sources of randomness. Hence, it differs from the approach described in Sect. 4, where only one source of uncertainty has been involved (i.e. the randomly generated vector error terms $\boldsymbol{\varepsilon}_t$). It should be highlighted that the proposed configuration of scenario generation can be replaced by any suitable alternative or can be further extended (see Sect. 4 for more details about incorporating discounting factors or interest rates).

Figures 5 and 6 display simulation results. Figure 5 clearly delivers analogous outputs as Fig. 3. However, the mean of forecasts is obviously more optimistic than those in Fig. 3. The empirical probability that the total non-life technical provision *TPT* in Q4 2017 will be greater than the two-thirds of the Q3 2015 level is 99.98%, which is comparable with Scenario 2 analysed in Sect. 4. Table 8 reports Value at Risk (*VaR*) and Expected Shortfall (*ES*) at the formal confidence level $\alpha = 0.5\%$ (see the discussion in Sect. 4). These values could be used as benchmarks for particular non-life technical provisions observed globally in the Czech insurance market.

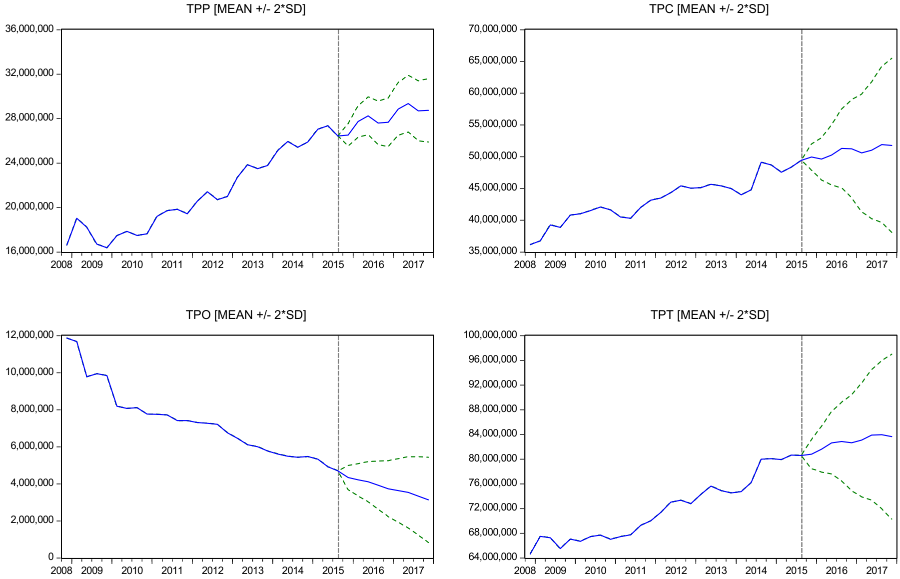

Fig. 5 The results of prognosing the particular non-life technical provisions: scenario analysis with randomly generated strictly exogenous variables (source: authors by EViews 8.0)

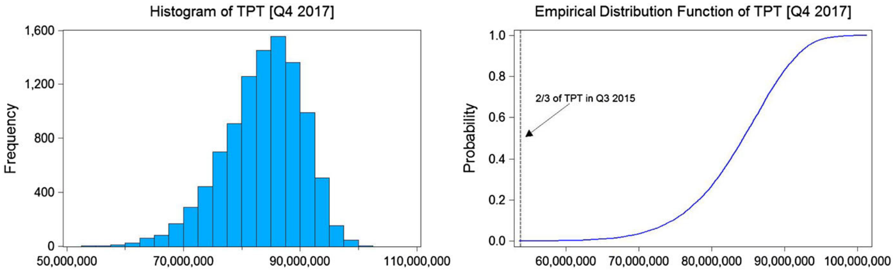

Fig. 6 The results of prognosing the total non-life technical provisions *TPT* for Q4 2017: scenario analysis with randomly generated strictly exogenous variables (source: authors by EViews 8.0)

Table 8 Value at Risk and Expected Shortfall measures (the confidence level $\alpha = 0.005$) for the predicted total non-life technical provisions *TPT* in Q4 2017: scenario analysis with randomly generated strictly exogenous variables

TPT Q4 2017	Random scenario
VaR	63076944
ES	61174135

Source: authors (by EViews 8.0)

6 Conclusions

The paper presented the complex econometric model of the key non-life technical provisions (and other important actuarial variables) based on the Czech non-life insurance market data. In particular, the econometric system of dynamic linear simultaneous equations was applied in order to describe economic-actuarial relationships within selected variables quarterly published in summary balance sheets of the Czech non-life insurers. The estimated modelling scheme was statistically verified and interpreted. Furthermore, we provided the scenario analysis by using the suggested modelling framework. It might be used e.g. for financial planning purposes, for stress testing, or for testing the sufficiency of the non-life technical provisions.

To illustrate the main idea of the discussed scenario testing, three underlying (extremely) unfavourable scenarios based on the prescribed strictly exogenous variables were investigated in greater detail. The simulation results derived by the residual bootstrap corresponded to pragmatic anticipations (i.e. the decreasing tendency of all the particular non-life technical provisions). Consequently, one identified that the number of existing non-life insurance contracts and the number of reported claims influenced the total non-life technical provisions substantially. Value at Risk and Expected Shortfall were calculated as measures of the sufficiency of the total non-life technical provisions and an instrument for adjustment of solvency capital requirements. Similar analysis was performed also for randomly generated strictly exogenous variables. The results corresponded to rational expectations.

One can summarize the introduced modelling scheme (1) in following points: (a) the simultaneous equations model (1) follows the typical econometric modelling framework (i.e. the dynamic system of linear simultaneous econometric equations), (b) it includes key technical provisions relevant from the perspective of non-life insurance, and it describes various economic-actuarial interactions among them, and (c) the numerical results confirm the model adequacy.

Acknowledgements This work was supported by the grant GA P402/12/G097. The authors thank to the Prudential Supervision Division of the Czech National Bank for the aggregated data used in the paper.

References

1. Baranoff EG, Papadopoulos S, Sager TW (2007) Capital and risk revisited: a structural equation model approach for life insurers. J Risk Insurance 74(3):653–681
2. Cipra T (1998) Econometric analysis of cash-flows in a life insurance company. Pojistné rozpravy 1998(3):62–72
3. Cipra T (2010) Financial and insurance formulas. Physica-Verlag/Springer, New York
4. Cipra T (2013) Financial econometrics, 2nd edn. Ekopress, Prague (in Czech)
5. Dahms R (2012) Linear stochastic reserving methods. ASTIN. Bulletin 42:1–34
6. EIOPA (2014) Technical specifications for the Solvency II preparatory phase (Part I). https://eiopa.europa.eu/regulation-supervision/insurance/solvency-ii/solvency-ii-technical-specifications
7. Feilmeier M, Junker M (1982) Bestandsorientierte Bilanzprojektion eines Lebensversicherungsunternehmens. Blätter Deutsche Gesellschaft für Versicherungsmathematik 15(3):287–306
8. Greene WH (2003) Econometric analysis. Prentice Hall, New Jersey
9. Hendrych R (2011) Econometric system of simultaneous equations in life insurance. In: Dlouhý M, Skočdopolová V (eds) Proceedings of the 29th international conference on mathematical methods in economics. Professional Publishing, Prague, pp 236–241
10. Hendrych R, Cipra T (2015) Econometric model of the Czech life insurance market. Prague Economic Papers 24(2):173–191
11. Hürlimann W (2009) Credible loss ratio claims reserves: the Benktander, Neuhaus and Mack methods revisited. ASTIN Bulletin 39:81–99
12. Lütkepohl H (2005) New introduction to multiple time series analysis. Springer, New York