Working Paper Series 12/2020

A Top-down Stress-testing Framework for the Nonfinancial Corporate Sector

Vojtěch Siuda

Czech National Bank —— Working Paper Series —— 12/2020

CZECH NATIONAL BANK

Reviewed by: Jiří Panoš (European Central Bank)

Petr Polák (Czech National Bank)

Project Coordinator: Zuzana Rakovská

A Top-down Stress-testing Framework for the Nonfinancial Corporate Sector

Vojtěch Siuda*

Abstract

This paper provides a framework for conducting simulations and stress testing in the non-financial corporations sector. It relies on national accounting and uses a set of input-output tables to track the propagation of shocks between parts of the sector while staying fully consistent with the big picture framed by the core forecasting model and the underlying scenario. The simulation framework allows standard macroeconomic developments to be captured, but one-off measures such as government wage and salary compensation and loan moratoria can also be easily implemented. The main output of the simulation is a set of industry-level performance and profitability variables. These variables can be used for various types of analysis, such as credit risk modelling and profitability and liquidity analysis. Some of them – such as the forecasting of portfolio default rates via learning process – are shown in the paper. The historical default rate estimates obtained are accurate and economically sensible for the majority of industries and exhibit a high degree of reliability even under very severe economic conditions. Given its national accounting framework and its level of detail, the model can be used to support decision-making processes and to evaluate the effects of existing or planned economic policies. Two different scenarios are considered to demonstrate the benefits of the proposed approach.

Abstrakt

Tento článek představuje rámec pro provádění simulací a zátěžových testů v sektoru nefinančních podniků. Ten vychází z národních účtů a využívá tzv. tabulky input-output ke sledování šíření šoků mezi jednotlivými součástmi sektoru. Je přitom plně konzistentní s širšími souvislostmi danými jádrovým predikčním modelem a příslušným scénářem. Simulační rámec umožňuje zachytit běžný makroekonomický vývoj, ale lze do něj snadno začlenit také jednorázová opatření, jako jsou vládní kompenzace mezd a platů či úvěrová moratoria. Hlavním výstupem simulace je sada proměnných popisujících výkon a ziskovost jednotlivých odvětví. Tyto proměnné je možné využít pro různé druhy analýz, například modelování úvěrového rizika či analýzy ziskovosti a likvidity. Některé z nich – například prognózování míry defaultu portfolia s využitím algoritmu strojového učení – jsou popsány v článku. Získané historické odhady míry defaultu jsou pro většinu odvětví přesné a ekonomicky smysluplné, přičemž vykazují vysokou spolehlivost i za velmi náročných ekonomických podmínek. Vzhledem k využití národních účtů a své podrobnosti může být model využit k podpoře rozhodovacího procesu a vyhodnocení dopadů stávajících nebo zamýšlených hospodářských politik. Výhody popsaného přístupu článek ukazuje na dvou rozdílných scénářích.

JEL Codes: G01, G32, H63.

Keywords: Credit default, default rate forecast, economic shock propagation, input-output tables.

---

* Vojtěch Siuda, Czech National Bank – Financial Stability Department, and University of Economics in Prague – Faculty of Finance and Accounting (vojtech.siuda@cnb.cz).

The research was supported by Czech Science Foundation Grant 18-05244S Innovation Approaches to Credit Risk Management, and by University of Economics IGA F1/46/2019.The views expressed in this paper are those of the author and do not necessarily reflect those of the Czech National Bank. I am very grateful to Miroslav Plašil for many inspiring comments and inputs for the scenario construction. I would also like to thank Adam Maryniok for helpful comments and suggestions. Nevertheless, none of the preceding is responsible for any errors and omissions contained in this paper.

1. Introduction

Given its share in gross value added, the non-financial corporations sector forms the backbone of all economies. However, firms often operate with significant leverage and account for a large part of total credit in the economy and are thus also a potential source of systemic risk. Therefore, it is necessary to understand the dynamics of shock propagation mechanisms in the economic system, regularly monitor potential sources of risks to firms in various industries, and ultimately quantify the associated credit losses of the financial sector. This can encourage the adoption of suitable measures for reducing systemic risk and restricting the effects of potential adverse shocks spreading through the economy.

Technological progress and growing international trade with decreasing taxes and levies have caused both the domestic and international interconnectedness of economies to increase dramatically during the 20th and the first two decades of the 21st century. The global economy experienced a strong negative demand shock during the Great Recession in the late 2000s. This spilled over very quickly from the US housing market and financial sector to the real economy and to other countries. At present, during the Covid-19 pandemic, the global economy is experiencing an even stronger and probably deeper negative shock from both the demand and supply sides. The need for tools capturing complex economic systems with both inter- and intra-sectoral links and providing an assessment of potential risks to the financial system becomes more evident with each new crisis.

This paper proposes such a tool. The presented framework facilitates the assessment of potential systemic risks stemming from corporate loan portfolios and quantifies defaults under various macroeconomic scenarios. The model can thus support a broad range of policy decisions, both longer-term and urgent. Our ambition was to create a robust tool with strong economic intuition and interpretation rather than a complicated black-box model with a high number of parameters, a perfect historical fit and potentially a low ability to generalize.

The two main contributions of the paper are the following. The first is the construction of a conceptual framework for simulating the propagation of shocks across the divisions of the economy given a predefined scenario. The second contribution lies in the identification of a varying set of determinants of sectional default rates, where the determinants include industry-specific indicators of economic performance and profitability (Section 4.1.). The combination of these two contributions opens up the possibility of forecasting default rates based on the economic performance of individual divisions and sections of non-financial corporations. To the best of our knowledge, the proposed approach is new in the literature and offers accurate estimates of historical default rates and reliable scenario-based projections.

The rest of the paper proceeds as follows. First, some relevant literature on modelling aggregate credit losses and its main findings are provided. The third section describes the data and the stress-testing methodology, consisting of economic simulation via input-output tables and a machine learning algorithm for default rate modelling. The next section presents the main outputs of the stress-testing exercise under two different macroeconomic scenarios, including default rates broken down by the sections of the economy (the NACE rev. 2 taxonomy is used throughout this paper, which defines 21 *sections* of economy on level 1 – letters A-U – and these are further divided into

88 *divisions* on level 2 – 2 digit code). Some additional outcomes which can serve as inputs for policy discussion are presented as well.

2. Related Literature

Stress testing of the banking sector's solvency became a common regulatory tool in the late 1990s[1] and developed apace after the global financial crisis, when stress tests also began to be used for assessing the resilience of other sectors. While a wide range of literature focuses on interconnectedness and systemic risk within the financial sector, a smaller volume has examined shocks spreading through the real economy and cascades of corporate defaults.

Stress testing in the corporate sector is commonly carried out on the aggregate level using a stylized sensitivity analysis often calibrated on static historical data. However, this approach hardly accounts for the relations between, and interactions of, the divisions of economic activity, which is a key feature of shock propagation. The macroeconomic determinants of the corporate default rate on the aggregate level were analysed by Karasulu and Jones (2006), who tested their ability to predict the impact of the crisis in 1997 on the Korean corporate sector. The authors found satisfactory predictive potential for corporate distress using a stylized sensitivity analysis of the aggregate corporate balance sheet. A similar methodology can be found in Feyen et al. (2017) and Klein (2016).

A shift to section-specific default rates and PD modelling was proposed by Virolainen (2004). His results show a significant relationship between default rates in some sections of the economy and aggregate whole-economy variables in Finland. A similar approach was pursued by Castrén et al. (2009), who used a global VAR model and expected default frequencies of individual firms to analyse European corporate sector default probabilities in an environment of macroeconomic and macrofinancial shocks. The identified significance of GDP growth, the exchange rate and stock and oil prices for default probabilities is in line with the rest of the literature. The effects of macroeconomic factors on corporate default rates using individual-firm data were analysed by Figlewski et al. (2012). They modelled rating migration with default intensity models and showed robust results under different economic conditions. However, this approach seems difficult to apply for stress-testing purposes and to intersectoral relations.

A different approach to modelling corporate defaults on the aggregate level was proposed by Drehmann (2005), who added fundamental factors (macroeconomic and market) to the structural Merton risk model. These risk factors exhibited a non-linearly increasing impact with rising shock severity. Cipollini and Missaglia (2008) went into even greater detail in their reduced-form dynamic factor model. They obtained encouraging results for aggregate divisional default rates. However, the applicability of their model for practical forward-looking policy-making may be limited due to its low flexibility in mapping between the macroeconomic scenario and economic performance on the level of individual divisions.

A number of studies strive to explain and estimate the effects of loan defaults on banks' balance sheets. The determinants of loan loss provisions and non-performing loan ratios were analysed, for

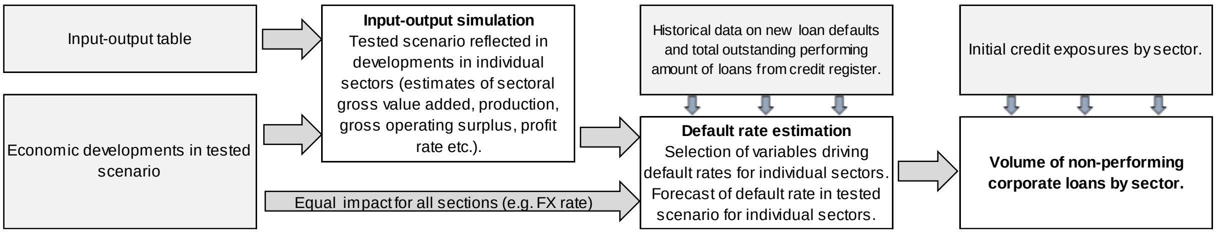
[1] Stress testing was recognized as method for market risk evaluation in the BCBS Market Risk Amendment in 1996. Stress tests were first used by the International Monetary Fund and World Bank in 1999 as part of the Financial Sector Assessment Program.

example, by Louzis et al. (2012), Radivojevic and Jovovic (2017) and Frait and Komarkova (2012). Contrary to these studies, this paper focuses on the quality of a loan portfolio associated with as yet unmaterialized risks. For this reason, our preferred measure of credit risk is the default rate, which should provide a more forward-looking indication of potential systemic stress in the future.

3. Data and Methodology

Our methodology can be divided into two main steps. First, a scenario based on aggregate economic numbers is run through an input-output simulation to obtain the performance of the individual divisions and sections of the economy. These section-specific performance variables are grouped with selected aggregate variables and form the basis for the second step, which consists in a default rate learning process. Figure 1 shows the simplified logic of the process.

Figure 1: Basic Logic of NFCs Stress-testing Framework



3.1 Input-Output Simulation

Input-output analysis was chosen as a main vehicle for modelling shock transmission between the divisions of the economy. This approach is preferable to computable general equilibrium models (see, for example, Dixon and Jorgenson, 2013), as the goal is not to predict economic developments themselves (which are given by the scenario), but rather to model the propagation of shocks and the interaction of economic divisions and their performance as implied by the underlying macroeconomic scenario. Input-output analysis is quite data demanding in terms of the necessary inputs, but this is more than outweighed by clear interpretation and practical applicability to real-world problems, which may help answer relevant policy questions. For the underlying macroeconomic scenario, one can take advantage of predictions made by government ministries, central banks or international institutions and use the outputs of their complex models. Economic scenarios can also be created independently, as shown in Section 4.3 of this paper. Designing ad hoc scenarios opens up space for a much wider range of analyses of possible economic developments that are hard to capture by the model families used currently. However, this approach also bears a higher risk of possible errors and inherent inconsistencies of variables and requires a deep knowledge of economic theory.

While the principles and objectives of financial accounting are generally known, common knowledge of national accounting is at a much lower level. National accounting relies on double-entry accounting of a nation's economic activity (on the resources side and the uses side). Since the input-output framework requires at least a basic knowledge of national accounting and input-output table construction, we briefly introduce several relations and identities which will be used throughout this paper. The starting point for this analysis is the production account, which tracks all goods and services entering the economy in a defined period. The resources side represents the sum of production and imports and price conversion from basic to purchasers' prices,[2] and the use side consists of intermediate consumption and final demand, which can be divided into households' consumption (C), government consumption (G), investment (I) and exports (Ex). Intermediate consumption tells us about the value of the goods and services consumed in the production process and, in the input-output table, describes the linkages between individual economic industries. The difference between production and intermediate consumption forms gross value added. Gross value added can also be viewed from an income perspective, as it equals the sum of compensation of employees, net taxes on production and gross operating surplus of firms. Finally, deducting consumption of fixed capital from gross operating surplus gives the net operating surplus. The input-output table used in this paper contains this information for 88 divisions.

The propagation of an economic shock across divisions builds on the well-known work of Leontief (1936). Consider an economy with S industries producing S products. Industries are connected via known input-output links and adjust their production in an infinite self-loop in each discrete step. Products produced or imported by industries are all consumed, stored in inventories or exported. In other words, there exists an identity where production at purchasers' prices and imports (together supply) equal intermediate consumption and final demand (together demand), which corresponds with the information in the production account.

To solve the open[3] model problem, let us denote:

$X = (x_1, x_2, \dots, x_S)^T, x_i > 0$ as a vector of supply of industries,

$A = \begin{pmatrix} \frac{a_{1,1}}{x_1} & \cdots & \frac{a_{S,1}}{x_1} \\ \vdots & \ddots & \vdots \\ \frac{a_{S,1}}{x_S} & \cdots & \frac{a_{S,S}}{x_S} \end{pmatrix}$ as a technology matrix,

$D = (d_1, d_2, \dots, d_S)^T, d_i \geq 0$ as a vector of final demand,

where all these variables are taken from the input-output matrix, and the technology matrix consists of the ratios of the intermediate consumption of segment i for the production of segment j to the supply of segment i,

The standard Leontief model defines supply as a linear combination of the technology matrix and final demand:

$$X = (I - A)^{-1}D = LD \quad (1)$$

where I denotes a unit matrix ($S \times S$) and L is known as the Leontief inverse matrix ($S \times S$). We can also rewrite the matrix for increments as:

$$\Delta X = L\Delta D \quad (2)$$

[2] This consists of taxes on products and subsidies on products, trade margins and transport margins.

[3] One can distinguish between two models according to the ultimate goal of the analysis. In an open model there exists a consumption vector $D = (d_1, d_2, .., d_S)^T$ for which at least one element d_i is positive. The problem solved is the production level if external demand is given. A closed model assumes that all production is consumed by industries (each element of final demand equals 0) and the problem solved is the relative price of each product (Hohn, 2004).

This notation implies that a shock to the final demand of a single segment i of size Δd_i will change the total supply by $\Delta d_i L^{(i)}$, where $L^{(i)}$ is the i-th column of matrix L.

The main criticism of the Leontief model relates to its static character. From the definition above, it is apparent that the input coefficients of production are constant over time. However, this hardly holds in practice. Empirical evidence shows that the coefficients evolve over time in line with structural changes in the economy and technological progress. Moreover, they can be expected to change more significantly during periods of strong distress. To overcome this limitation, we borrow from the approach of Alaniste Contreras and Fagiolo (2014) and adjust one of their models (the third one).

Suppose a known definition for X, A, D and L and a new vector of shocks to final demand $Q = (q_1, q_2, \dots, q_S)^T, q_i \geq 0$. A segment k hit by a shock q_k has less to produce and less to supply to other segments, which changes the whole intermediate consumption matrix and thus the technology matrix A. This leads to a new technology matrix A', where each element in the k-th row and k-th column has been updated by q_k to:

$$a'_{k,j} = q_k a_{k,j} \tag{3}$$

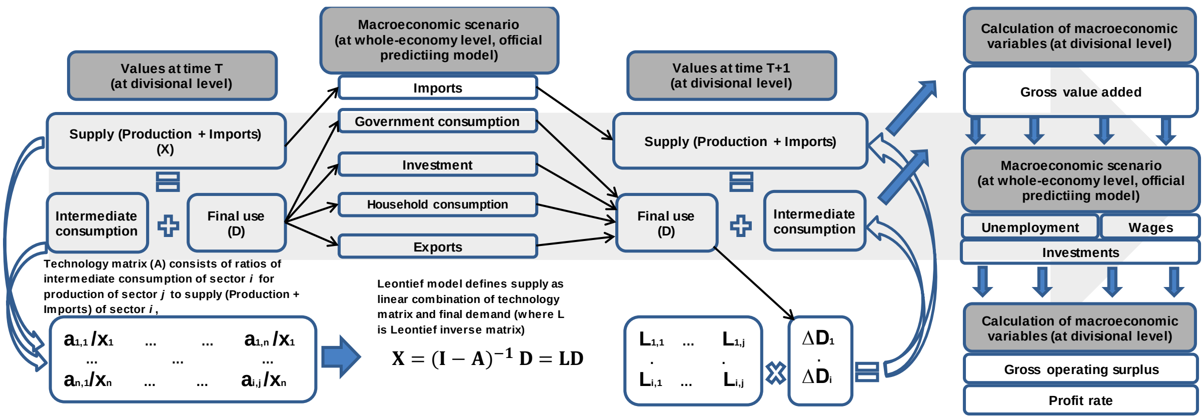
$$a'_{i,k} = q_k a_{i,k} \tag{4}$$

where j is any segment that uses a commodity produced by segment k as an input and i is any segment from which k buys inputs. Finally, we obtain the new production increments vector:

$$\Delta X = (I - A')^{-1} \Delta D = L' \Delta D \tag{5}$$

This mechanism assumes accurate foresight of firms about economic activity. It is a self-fulfilling process and adjusts the coefficients in the technology matrix, which adapts immediately and fully to the new level of final demand. With a defined change of final demand, one can identify the level of supply (and, with known imports, also the level of production), intermediate consumption and gross value added of each segment of the economy in each discrete step. The simplified macroeconomic simulation process is shown in Figure 2.

Figure 2: Simplified One Step of Macrosimulation



The speed of adaptation of intersectoral relations has risen strongly over the past few decades and, in our opinion, the frequency of possible changes and substitutions in the supply chain is higher than the frequency of publication of input-output tables. To capture this higher frequency, the starting input-output table was recalculated from yearly to seasonally adjusted quarterly frequency for the purposes of this simulation.

One of the advantages of the chosen model is that it offers some flexibility in the interpretation of aggregate scenario numbers. Usually, the scenario is defined in the big picture of the whole economy. Its impact can be projected into the individual divisions in two ways – as a uniform impact on all divisions, or as a differentiated impact. With a uniform impact, the final demand of each division is hit equally by the aggregate numbers.[4] Under the differentiated approach, the shock impacts on individual divisions differently (there is no fixed matching of the impact to the final demand components for the individual divisions), which means that judgement about the specific economic story can be assigned to the underlying aggregate scenario, leaving more space for calibration of the shock with greater detail on the chosen segments of economic activity.[5] This opens up the possibility of increasing the stress on some segments and reducing that on others.

3.2 Default Rate Estimation

Based on the set of quantitative outputs obtained from the input-output framework, we need to select the optimal set of variables for predicting default rates in individual industries. Agarwal and Taffler (2008) claim that the decision of banks to extend loans relies on the submission of financial and accounting data from applicants. Accounting information and a set of accounting ratios are used by a bank to derive a firm's credit quality and related credit risk. The variables entering the default rate model in our approach draw on similar information converted into aggregate performance ratios or close approximates thereof (Table 1). To reduce excessive noise in the data, we only work with default rates on NACE level 1 (the divisional results from the input-output simulation were merged into sectional ones). A division-level breakdown would be possible for larger divisions, for example from manufacturing industry, but a broader view was necessary to strike a balance between systematic patterns and pure noise-fitting.

Table 1: Accounting Variables and Their Macroeconomic Equivalents

Macro variables	Accounting variables
$\frac{Net\ operating\ surplus_t}{Production_t}$	Corresponds to profits after depreciation and taxes normalized by sales (Net earnings/Sales)
$\frac{Gross\ operating\ surplus_t}{Gross\ value\ added_t}$	(EBITDA/Gross profit)
$\frac{Gross\ operating\ surplus_t}{Production_t}$	Corresponds to gross operating profit to sales (EBITDA/Sales)
$\frac{Gross\ operating\ surplus}{Average\ interest\ rate * volume\ of\ credit}$	Close to interest coverage ratio (EBITDA/Interest expenses)
$\frac{Gross\ value\ added_t}{Gross\ value\ added_{t-1}}$	Corresponds to gross profit dynamics

[4] For example, when aggregate exports rise by 3%, the exports of each division increase by 3%. However, this does not mean that the shock will have identical impacts on the individual divisions, because each division has a different composition of final demand.

[5] That is, with respect to the total impact of the shock, the structural characteristics of the economy, and the implicit economic rules (respecting, for example, the import intensities of exports and investment).

The relation between macro variables and accounting variables can be described as follows. Production contains information on all goods produced and services provided across the economy. In the production process, the inputs of labour, capital and goods and services are used to produce outputs of goods and services (Lequiller and Blades, 2014). In individual firms' accounting, production is the mirror of total sales. Production forms the basis for gross value added. The products used in production are represented by intermediate consumption. Deducting intermediate consumption from production eliminates double counting in the production process and produces gross value added. At the firm level, gross value added measures the net value of the firm's operations (before deduction of taxes, wages and depreciation). Gross value added on both the aggregate and firm level can also be expressed as the sum of gross operating surplus/gross profit (GOS), labour costs and net taxes (taxes minus subsidies). By deducting capital depreciation from gross operating surplus one obtains the net operating surplus (NOS), which is the highest level of detail that can be achieved on the input-output analysis level. Finally, the exchange rate and property prices were added to the list of potential drivers of default rates, as empirical evidence suggests that these macro variables strongly affect the financial performance of some sections of nonfinancial corporates (Blackwood, 2017, describes the link between property prices and firm performance; for the exchange rate, see, for example, Baum et al., 2001).[6]

The initial list consists of 20 variables, a mix of section-specific ones (based on Table 1) and aggregate ones and their transformations and lags.[7] Table 2 summarizes the set of all variables initially entering the default rate modelling, including the logical sign restrictions imposed on the coefficients' values. Profit variables enter in the form of either ratios or growth rates (year-on-year change), while gross value added and the exchange rate only enter in growth rates, and interest paid is only defined as a ratio.

Standard reduced-form credit risk modelling in both default-only mode (only default risk quantification) and migration mode (mark-to-market losses also taken into account) differs from the presented aggregate macroeconomic default modelling in its main interpretation. Quantitative credit risk estimation aims to maximize the lender's risk-adjusted rate of return while maintaining acceptable risk exposure parameters (BCBS, 2000). The main purpose of credit risk modelling is to make the best probabilistic assessment of the likelihood of default for each single loan. It usually uses a very large historical dataset which always contains a binary dependent default variable, a set of possible explanatory variables and a non-linear model specification. The outcome of this process is the loan-specific probability of default (in default mode). The dichotomy between the standard credit risk model and the model presented here lies in the very nature of these approaches. The presented approach does not work with a dependent binary variable but with the aggregate continuous default rate and is designed to identify potential risks to the stability of the financial sector under severe conditions (tail events) and to help answer crucial policy questions. A model with clear economic interpretation is more suitable for this purpose. A high degree of interpretability is achieved by including only economically justifiable variables and a set of logical restrictions on

[6] A large share of FX loans is specific for Czech non-financial corporate sector (near 33%). These were firstly splitted according to domestic currency and foreign currency in a learning process, but the overall fit of historical default rates was a little worse than for loan stocks aggregated by currency.

[7] Considering the length of time series and the decreasing importance of lagged variables with increasing lag (Simons and Rolwes, 2008), only one-year-lagged variables were included in the list.

the parameter values. Predictions of the aggregate amount of loan defaults in each section of the economy in the given scenario are then the final outcome of the model.

Table 2: List of All Section-specific and Aggregate Variables Initially Entering Default Rate Learning Algorithm

	Variable	**Abbreviation**	**Restriction**
Dependent variable	Default rate$_t$ => default rate at time t	DF_t	
Section-specific explanatory variables	Constant	Const.	
	Default rate$_{t-1}$ => default rate at time t-1 (AR1 process)	DF_{t-1}	>0
	Net operating surplus$_t$/Production$_t$	$(NOS/P)_t$	<0
	Net operating surplus$_{t-1}$/Production$_{t-1}$	$(NOS/P)_{t-1}$	<0
	Gross operating surplus$_t$/Gross value added$_t$	$(GOS/GVA)_t$	<0
	Gross operating surplus$_{t-1}$/Gross value added$_{t-1}$	$(GOS/GVA)_{t-1}$	<0
	Gross operating surplus$_t$/Production$_t$	$(GOS/P)_t$	<0
	Gross operating surplus$_{t-1}$/Production$_{t-1}$	$(GOS/P)_{t-1}$	<0
	(Net operating surplus$_t$ – Net operating surplus$_{t-1}$)/ Production$_{t-1}$	$\Delta(NOS/P)_t$	<0
	(Net operating surplus$_{t-1}$ – Net operating surplus$_{t-2}$)/ Production$_{t-2}$	$\Delta(NOS/P)_{t-1}$	<0
	(Gross operating surplus$_t$ – Gross operating surplus$_{t-1}$)/ Gross value added$_{t-1}$	$\Delta(GOS/GVA)_t$	<0
	(Gross operating surplus$_{t-1}$ – Gross operating surplus$_{t-2}$)/ Gross value added$_{t-2}$	$\Delta(GOS/GVA)_{t-1}$	<0
	(Gross operating surplus$_t$ – Net operating surplus$_{t-1}$)/ Production$_{t-1}$	$\Delta(GOS/P)_t$	<0
	(Gross operating surplus$_{t-1}$ – Net operating surplus$_{t-2}$)/ Production$_{t-2}$	$\Delta(GOS/P)_{t-1}$	<0
	Gross value added$_t$/Gross value added$_{t-1}$	ΔGVA_t	<0
	Gross value added$_{t-1}$/Gross value added$_{t-2}$	ΔGVA_{t-1}	<0
	Gross operating surplus$_t$/Interest paid$_t$	ICR_t	<0
	Gross operating surplus$_{t-1}$/Interest paid$_{t-1}$	ICR_{t-1}	<0
Aggregate explanatory variables	Property prices$_t$/Property prices$_{t-1}$	ΔPP_t	
	Property prices$_{t-1}$Property prices$_{t-2}$	ΔPP_{t-1}	
	FX_t/FX_{t-1}*	ΔFX_t	
	*Nominal FX rate is in direct quotation => higher value means relative depreciation		

There are several basic approaches to variable selection. Importantly, many more complex approaches already exist for optimal model and variable selection in default rate forecasting (extensive work on model selection is presented, for example, by Panoš and Polák, 2020). However, the approach presented in this paper is not focused solely on forecasting default rates, but rather presents an overall conceptual framework. In line with this, a less computationally intensive and complex but more interpretable approach was chosen. In our view, among the less complex approaches, stepwise selection remains the most common credit risk method despite all the related issues (as summarized, for example, by Harrell, 2001, and Flom, 2018). Nevertheless, instead of stepwise methods, we employed LASSO estimator (Tibshirani, 1996) to select optimal set of predictors and regularize the problem. This is a machine learning algorithm that shrinks the parameters of less significant variables to 0.

Consider a vector of observations $Y = (y_1, y_2, \dots, y_T)$ of length T and a $T \times P$ matrix of predictors $X = (x_{1,1}, \dots, x_{T,1}, \dots, x_{1,P}, \dots, x_{T,P})$. Each column of X (each predictor) is standardized to have zero mean and unit variance. Classical linear regression minimizes the following term:

$$min \sum_{t=1}^{T} \left(y_t - \beta_p X_{t,p}\right)^2 \tag{6}$$

where $\beta = (\beta_1 . \beta_2, \ldots, \beta_P)$ is the vector of regression coefficients. Adding a penalizing term in the form of the absolute value of the coefficients gives the LASSO estimator:

$$min \left(\sum_{t=1}^{T} \left(y_t - \beta_p X_{t,p}\right)^2 + \lambda \left|\beta_p\right| \right) \tag{7}$$

where λ is the tuning parameter that regulates the degree of shrinkage.[8] Generally, the superiority of regularization techniques over standard subset, forward, backward or stepwise selection neatly avoids overfitting bias. This problem tends to be even more significant in situations where the available time series are not long enough (as in the case of emerging or converging economies). Moreover, penalizing the loss function $(y_t - \beta_p X_{t,p})^2$ with the coefficient value also solves a problem in the presented approach where the defined explanatory variables tend to be correlated and shrink the less significant ones. The preference for LASSO penalization over ridge regression[9] in this model lies in clearer interpretation of the results. Moreover, the squared penalty term in ridge regression means that the larger the parameters, the more they are penalized, whereas LASSO penalizes them more linearly. This means that if there is a powerful predictor in our list of variables for forecasting, the predictor's effectiveness is shrunk more by the ridge than the LASSO. The coefficient values were restricted to respect the conventional logical interpretation (for example, rising profits cannot be a source of growth in default rates) to ensure robustness of the estimates. This feature will further improve the parameters' stability. The algorithm was performed in R with the glmnet package (Hastie and Qian, 2016). The optimal value of the tuning parameter λ for each section was optimized by cross-validation.[10]

3.3 Data

Input-output tables (available from the Czech Statistical Office) were used for the simulation. The simulation itself starts with the last known input-output table (the one for 2018 at the time of writing). Data on the components of gross value added (cost of labour, net taxes, consumption of fixed capital, operating surplus and mixed income) by each segment of economic activity[11] were obtained from the same source (historical input-output tables) and time series of those data were created. These time series were used for the learning process in modelling the historical default rates as described in Section 4.1. Property price data (available from the Czech Statistical Office) and EUR/CZK exchange rate and corporate loan rate data (available from the Czech National Bank) were also used for modelling the historical default rates.

[8] Obviously, lambda equalling zero returns a standard linear regression.

[9] Ridge regression penalizes the loss function with squared coefficients and none of the coefficients shrinks exactly to zero. A possible compromise between these two forms is the elastic net, which uses both types of penalization.

[10] The criterion for lambda optimization was root mean squared error (see also Figures A3 and A4 in the Appendix).

[11] Wages and salaries, employers' social contributions, taxes and subsidies on production, consumption of fixed capital and net operating surplus.

A Deep Recession scenario based on the Adverse Scenario published in *Risks to Financial Stability and Their Indicators 2019* by the Czech National Bank was used to simulate future developments. The variables used in the simulation and forecast are:

a) quarterly data on gross domestic product, including its components, and the partition of the net trade balance into imports and exports as an input for final demand and for calculating the level of production itself.

b) quarterly labour market variables, including nominal wage growth and the unemployment rate, for calculating the evolution of wages and salaries.

c) quarterly corporate credit growth, interest rates and property prices as implied by the economic scenario used in the Czech National Bank's satellite models, for calculating the profit variables entering the default rate estimation.

d) banks' total corporate credit for the individual divisions, including its maturity structure and interest rates, for calculating quarterly average interest and total debt payments for each division of economy.

Finally, data on new non-performing loans and total volumes of performing loans were taken from the Czech Republic's Central Credit Register managed by the Czech National Bank. These data were used to calculate historical default rates at the level of each section of the economy. Default rates are defined in a forward-looking manner as:

$$Default\ rate_{t,i} = \frac{NDL_{t+1,i}}{TOPA_{t,i}} \tag{8}$$

where $NDL_{t+1,i}$ is the amount of new loan defaults in period $t+1$ in section i and $TOPA_{t,i}$ is the total outstanding amount of performing loans at time t in section i.

4. Results for the Czech Republic

4.1 Historical Default Rate Estimates

The presented algorithm was trained on the 12M default rates of Czech corporates from 2004 to 2018. The parameter values and explained variance are reported in Table 3.

Interest payments (12 non-zero coefficients at time t and 11 at time t-1) were selected most frequently for the individual sections, followed by the net surplus dynamics (8 and 10) and economic performance as expressed by the dynamics of gross value added (9 and 8). In general, the individual sections are sensitive to the dynamics of profit rather than to the profit ratio levels themselves. This indicates that firms may not accumulate sufficient reserves in good times (when the ratios are higher) and default rates tend to rise very quickly when profits deteriorate.

Table 3: Coefficients Obtained for Individual Sections from Learning Algorithm

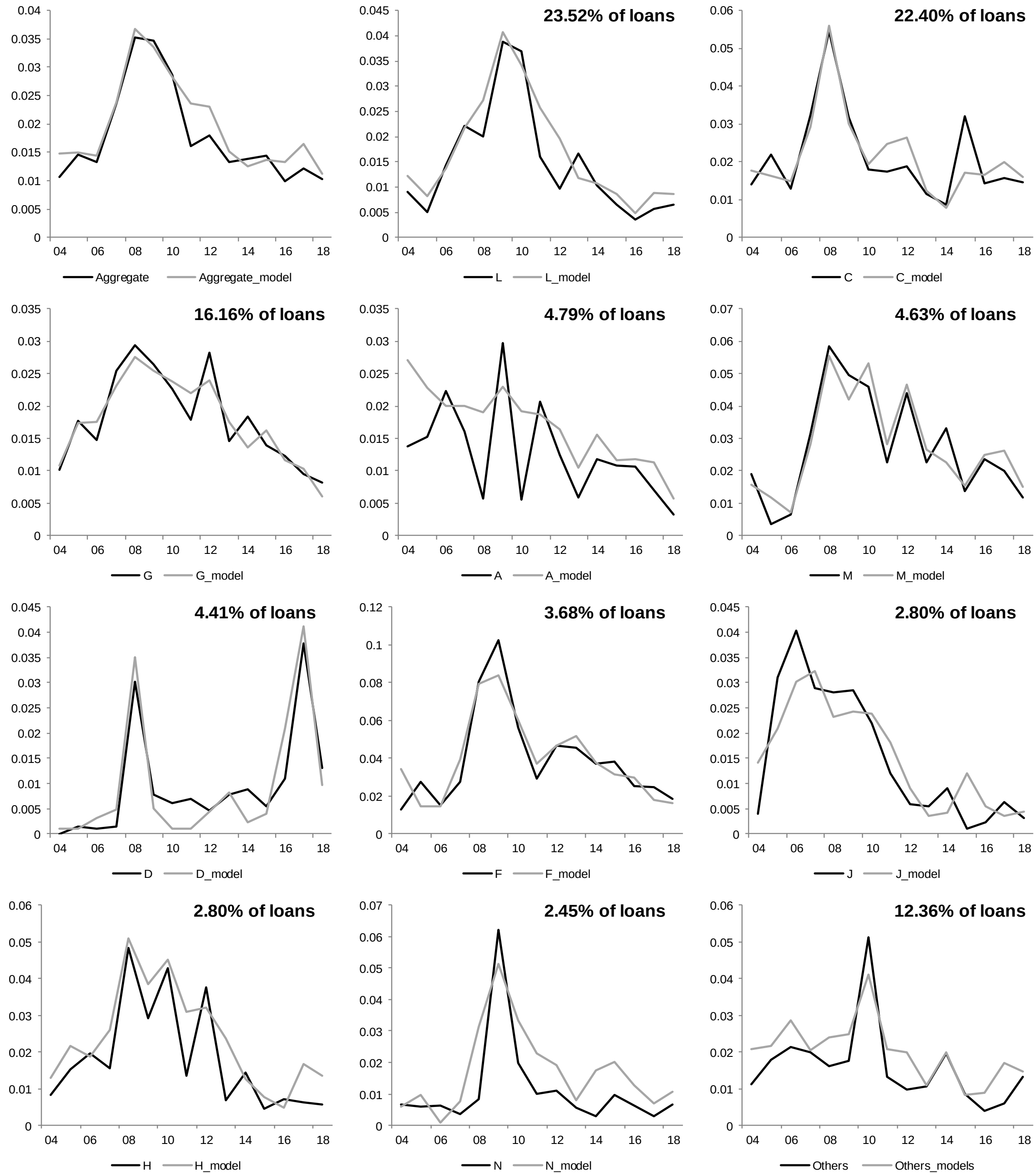
Variable/NACE	A	B	C	D	E	F	G	H	I	J
Const.	0.02	0.002	0.021	0.003	0.001	0.027	0.018	0.025	0.082	0.008
DF_{t-1}	0	0	0	0	0	0.009	0	0	0	0.007
$\Delta(NOS/P)_t$	-0.001	-0.001	-0.002	-0.012	0	-0.004	0	0	0	-0.001
$\Delta(NOS/P)_{t-1}$	-0.002	-0.014	0	-0.006	0	-0.013	-0.003	0	0	-0.002
$\Delta(GOP/GVA)_t$	0	0	0	-0.011	0	0	0	0	-0.001	0
$\Delta(GOP/GVA)_{t-1}$	0	0	-0.002	-0.001	0	0	0	0	0	0
$\Delta(GOP/P)_t$	0	-0.004	0	0	0	0	-0.001	-0.005	0	0
$\Delta(GOP/P)_{t-1}$	0	-0.009	0	-0.001	0	0	0	0	0	0
$(NOS/P)_t$	0	0	0	0	0	0	-0.001	0	0	0
$(NOS/P)_{t-1}$	-0.003	-0.004	0	0	0	0	0	-0.018	0	0
$(GOP/GVA)_t$	0	0	0	0	0	0	0	0	0	0
$(GOP/GVA)_{t-1}$	-0.001	0	0	0	0	0	0	0	0	0
$(GOP/P)_t$	-0.006	0	0	0	0	0	0	0	0	0
$(GOP/P)_{t-1}$	0	0	0	0	0	0	0	0	0	0
ΔGVA_t	-0.001	0	-0.007	0	0	0	-0.001	-0.003	0	0
ΔGVA_{t-1}	0	-0.004	-0.003	0	0	-0.006	-0.002	0	-0.016	0
ΔPP_t	0	0	0	0	0	0	0	0	0	0
ΔPP_{t-1}	0	0	0	0	0	-0.004	0	0	0	0
ΔFX_t	0	-0.001	0	0	0	0	0	0	0	0
ICR_t	-0.003	-0.002	0	0.004	0	0	0.001	-0.001	0	-0.006
ICR_{t-1}	0	0.002	0.004	0	0	0.016	0.004	0.016	0.009	0
Lambda	0.000035	0.000025	0.000148	0.000043	0.006	0.001023	0.000171	0.000102	0.008249	0.000924

Variable/NACE	K	L	M	N	O	P	Q	R	S	Unclass
Const.	0.001	0.016	0.028	0.019	0	0.038	0.017	0.037	0.037	0.007
DF_{t-1}	0	0	0	0	0	0	0	0	0	0
$\Delta(NOS/P)_t$	0	0	-0.002	0	0	0	0	0	-0.025	0
$\Delta(NOS/P)_{t-1}$	0	-0.001	0	0	0	0	-0.007	0	-0.006	-0.005
$\Delta(GOP/GVA)_t$	0	0	0	-0.011	0	0	0	0	0	0
$\Delta(GOP/GVA)_{t-1}$	0	0	-0.008	0	0	0	0	0	0	0
$\Delta(GOP/P)_t$	0	-0.004	0	0	0	0	0	0	-0.034	-0.005
$\Delta(GOP/P)_{t-1}$	0	-0.004	0	-0.005	0	0	0	0	-0.022	-0.01
$(NOS/P)_t$	0	0	0	0	0	0	0	0	0	0
$(NOS/P)_{t-1}$	0	0	0	0	0	0	0	0	0	0
$(GOP/GVA)_t$	0	0	0	0	0	0	0	0	0	0
$(GOP/GVA)_{t-1}$	0	0	0	0	0	0	0	0	0	0
$(GOP/P)_t$	0	0	0	0	0	0	0	0	0	0
$(GOP/P)_{t-1}$	0	0	0	0	0	0	0	0	0	0
ΔGVA_t	0	0	-0.009	0	0	-0.006	0	-0.005	0	-0.005
ΔGVA_{t-1}	0	0	0	-0.006	0	0	-0.001	-0.002	0	-0.002
ΔPP_t	0	-0.002	0	0	0	0	0	0	0	0
ΔPP_{t-1}	0	-0.006	0	0	0	0	0	0	0	0
ΔFX_t	0	0	0	0	0	0	0	0	0	0
ICR_t	0	0.004	-0.005	-0.006	0	0	0	-0.004	-0.006	-0.025
ICR_{t-1}	0	0.004	0	0.004	0	0	0	0.001	0.005	0.008
Lambda	0.001543	0.000066	0.000650	0.000112	0.000025	0.009704	0.000888	0.001932	0.000166	0.000042

Note: A – Agriculture, forestry and fishing, B – Mining and quarrying, C – Manufacturing, D – Electricity, gas, steam and air conditioning supply, E – Water supply, sewerage and waste management, F – Construction, G – Wholesale and retail trade, repair of motor vehicles and motorcycles, H – Transportation and storage, I – Accommodation and food service activities, J – Information and communication, K – Financial and insurance activities, L – Real estate activities, M – Professional, scientific and technical activities, N – Administrative and support service activities, O – Public administration and defence, compulsory social security, P – Education, Q – Human health and social work activities, R – Arts, entertainment and recreation, S – Other service activities, Unclass – Loans without reported NACE code, whole economy averages used for learning process. Estimates for sections A, H, L and N were performed on filtered series (Hodrick-Prescott with $\lambda = 0.5$) due to very low signal-to-noise ratio.

Source: CZSO, CNB, author's calculations

Figure 3: Historical Default Rate Estimates



Note: A – Agriculture, forestry and fishing, C – Manufacturing, D – Electricity, gas, steam and air conditioning supply, sewerage and waste management, F – Construction, G – Wholesale and retail trade, repair of motor vehicles and motorcycles, H – Transportation and storage, J – Information and communication, L – Real estate activities, M – Professional, scientific and technical activities, N – Administrative and support service activities. Due to low materiality, sections B – Mining and quarrying, E – Water supply, I – Accommodation and food service activities, K – Financial and insurance activities, O – Public administration and defence, compulsory social security, P – Education, Q – Human health and social work activities, R – Arts, entertainment and recreation and S – Other service activities were combined with unclassified loans as Others. Estimates for sections A, H, L and N were performed on HP ($\lambda = 0.5$) filtered series due to very low signal-to-noise ratio. Only real default rates and estimates are shown. Percentages in subplots represent share of individual sections in NFCs' total loan stock.

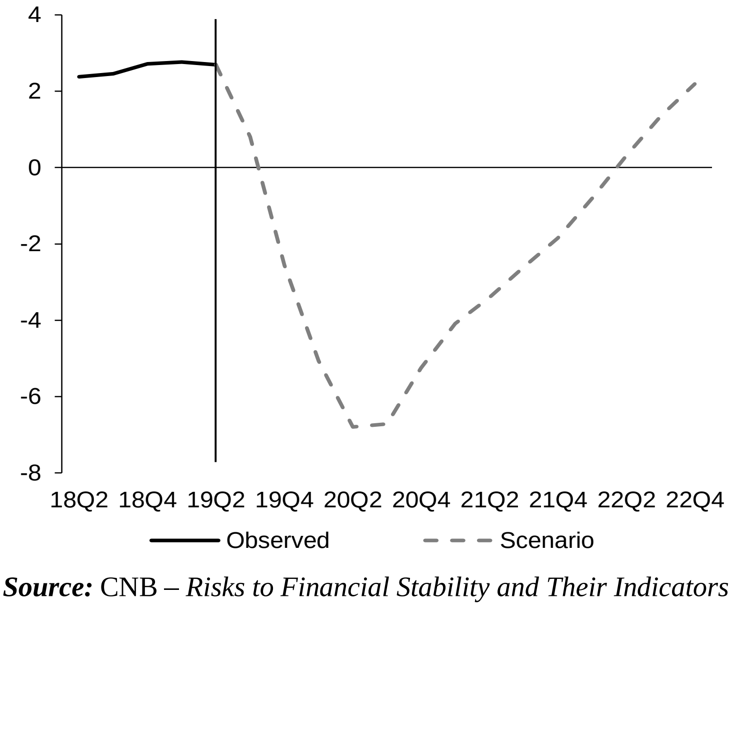
Source: CNB, author's calculations

Figure 3 shows the default rate estimates for 2004–2018. The aggregate default rate is calculated as a weighted average of the sectional default rates. Due to low materiality, high noise in the time series and concentrated portfolios, sections with a share of total loans of less than 2% (5.6% of the total outstanding performing amount) were merged with unclassified loans (6.8% of the total outstanding performing amount) in the Others category. The out-of-sample characteristics and performance, including the mean squared errors of the estimates based on the lambda value and the number of variables chosen, are shown in the Appendix (Figure A3). A plot showing the parameter values depending on lambda can also be found in the Appendix (Figure A4). Nevertheless, the presented approach is based not on a statistical model, but on a learning model where no assumptions were made about the underlying process. Given the nature of the learning algorithm, there is no need for other extensive model diagnostics and stability checks.

4.2 Scenario A – Deep Recession

The first presented scenario – Deep Recession – corresponds to the Adverse Scenario published by the CNB in *Risks to Financial Stability and Their Indicators* in December 2019. It assumes a strong demand shock originating in, and imported from, the foreign economy. This leads to a sharp economic slowdown in the Czech Republic over a three-year horizon. The falling external demand strongly affects export-oriented sections of the economy and causes the net trade balance to drop dramatically from a quarterly surplus of approximately CZK 100 billion to near zero in the third year of the scenario. Gross capital formation also experiences a strong shock (a 20% slump YoY at the peak of the crisis in real terms). In response to the decrease in investment, growth of corporate loans falls rapidly into negative figures. Private consumption shows greater resistance to the shock (falling by 3% YoY at the peak of the crisis in real terms), while government consumption acts countercyclically and supports the economy. The scenario also implies rising unemployment and a drop in nominal wages and salaries. The lower demand pushes down property prices (by 15% YoY at the peak of the crisis and 24% cumulatively during the scenario). With the credit spread widening, corporate interest rates rise. The CZK/EUR rate depreciates sharply (by about 14%) in the first year of the scenario, amid increased uncertainty and global risk aversion, and then stays almost unchanged until 2022. GDP growth is shown in Figure 4.

Figure 4: Real GDP Growth Projection in Deep Recession Scenario

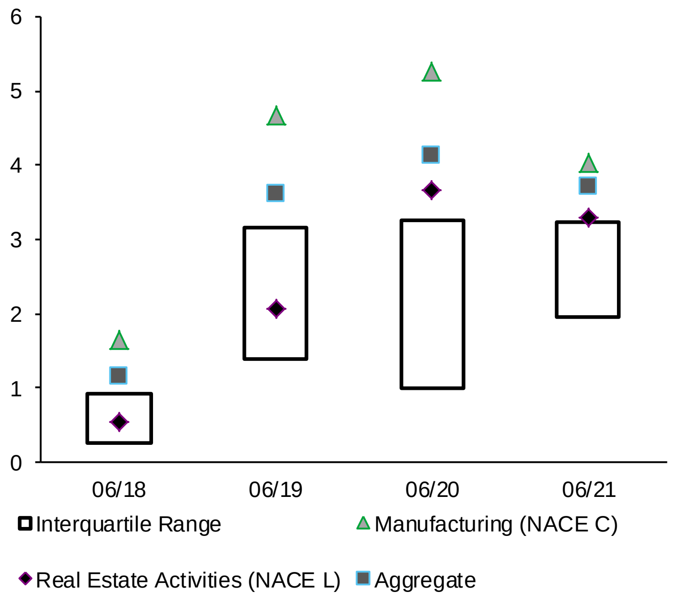


Source: CNB – *Risks to Financial Stability and Their Indicators*

For simplification, the shock to each component of GDP (final demand and imports) was proportionately divided into each division of the economy. Aggregate production in real terms falls to 95% of its maximum in the third year. Persistence in costs drives operating profits to decrease sharply (the impacts on selected sections and divisions are displayed in Figure A1). The presented scenario shows a significant rise in aggregate corporate defaults (of over 4% at the peak of the crisis), slightly exceeding the values observed during the Great Recession in the late 2000s. Despite a deeper and much longer recession than in 2009 and 2010, the forecasts show that non-financial corporates would not default significantly more. This may be explained by their lower relative indebtedness and lower debt service costs, indicating that they are in better financial health overall than they were in the late 2000s. The biggest contribution to the rise in the aggregate default rate is recorded by manufacturing industry, which is dependent on strongly affected international trade. Real estate activities also contribute notably to the aggregate default rate, owing to sharply falling property prices. Overall, the model shows that about 11.4% of performing loans will default over the three-year horizon of the stress simulation (Figure 5). The aggregate default rate is composed of the individual default rates of each division.

For transposition of the results in terms of the capital ratio, we need to know the LGD parameter and the risk weights. For simplification, a constant 50% LGD and static risk-weights were assumed.[12] Considering the total exposure of banks to the non-financial corporations sector (around CZK 1.5 billion), the losses run to CZK 87 billion, which corresponds to 3.42% of total risk-weighted assets.[13]

Figure 5: Default Rate Forecasts in Deep Recession Scenario (in %)



Note: 06/18 represents observed data.
Source: Observed data – CNB; forecasts – author's calculations

[12] It is not the goal of this paper to provide forecasts of LGD. This calculation serves only as an example of a possible outcome.

[13] 3Y default rate = 0.1144, LGD = 0.5, total risk-weighted assets = CZK 2,557 billion, NFC exposure = CZK 15.29 billion; 1,529*0.1144*0.5/2,557 = 0.0342.

4.3 Scenario B – Covid-19

As mentioned earlier, at the time of writing the global economy was being strongly affected by government measures to fight the spread of the Covid-19 pandemic. Economic activity was frozen with unprecedented speed. Many firms and whole divisions of the economy were forced to restrict their activities due to “anti-pandemic” measures, a lack of production inputs caused by supply-chain disruptions, and shortages of employees. Policy questions about the ways and means of supporting the economy arose immediately.

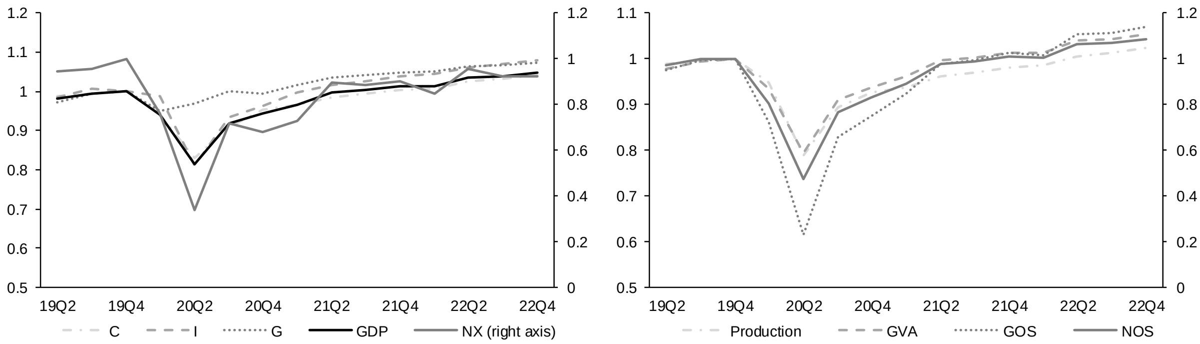
Given the rapid developments, one of the main problems during the Covid-19 crisis has been a lack of relevant, live information. Forecasting some of the possible macroeconomic impacts of the Covid-19 shock in the early stage of the crisis was our main motivation for creating a Covid-19 scenario for macroeconomic simulations and default rate forecasting. During March and early April 2020, a detailed macroeconomic scenario with a differentiated impact on the individual divisions of the economy was created for the Czech economy. The scenario was constructed mainly by collecting information from various data sources, such as news drilling, press agencies’ statements, information from ministries and governmental agencies, electricity consumption and other at least remotely relevant information sources. This information was used as a guide for expert judgement in quantifying the intensity of the crisis and also to set the initial controlling variables[14] for the computations in the input-output table. The scenario was focused on the 30 largest divisions of the economy according to their shares in gross value added (the “core”, accounting for around 80% of total gross value added). For the rest of the economy the averages of the core were used. For each division in the core, an assessment was made about each component of final demand and imports for every two weeks until the end of 2020. Given the supply as well as demand character of the shock (at least in the first stage), an additional primary assumption about the supply of each division in the core was made in the same way as for final demand. A condition was added to the supply estimation algorithm whereby the level of supply implied by the demand shock (eq. 2 and 3) was compared with the primary assumption about supply and the lower value was chosen. If the primary supply assumption was chosen, intermediate consumption and final demand were adjusted proportionately to respect the macroeconomic identity that supply equals demand. The logic of this condition is clear from the simplified R notation (Figure A2).

The presented scenario implied immediate negative GDP growth (-2%) in 2020Q1, followed by a massive drop (of 17.1% YoY) in 2020Q2 and then a gradual recovery of the economy. Cumulatively, GDP falls by about 8.4% in 2020 in the Covid-19 scenario. For 2021 and 2022, the scenario implies convergence to the pre-crisis composition of the economy (no structural change in the long term is assumed), which means that the more strongly affected divisions will grow more quickly and, after a sharper rebound in 2021, GDP growth converges to the steady state (around 3% in real terms). The Czech koruna depreciates by 10% against the euro in 2020Q1, in line with the high uncertainty on financial markets and capital flows to safe havens, then pulls back slowly and returns to the pre-crisis level at the scenario horizon. The drop in investment also strangles corporate credit growth, which is negative from 2020Q1 to 2021Q2. Growth in property prices slows but does

[14] For example, information about a drop in electricity consumption was captured in a macroeconomic simulation, and the final use of the divisions was calibrated to cause a corresponding drop in production in NACE category D – the energy sector.

not turn negative due to a sharp decrease in interest rates, which encourages purchasing activity on the housing market. The macroeconomic picture of the Covid-19 scenario is presented in Figure 6.

Figure 6: Covid-19 Scenario – Macroeconomic Description (Real Terms Index: 2019Q4 = 1)



Note: C stands for consumption, I for investment, G for government expenditure and NX for net exports.
Source: Author's calculations

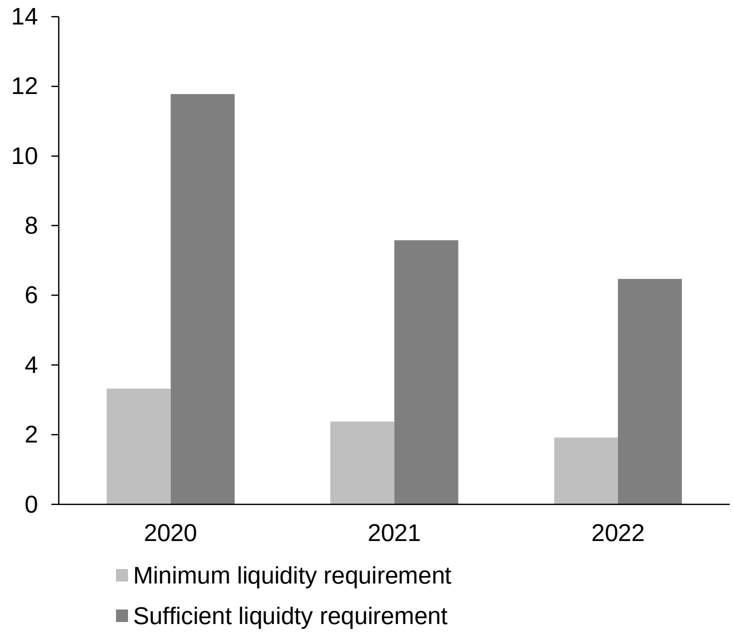
One of the most urgent policy questions was about the liquidity need in the non-financial corporations sector. Considering the prescribed lockdown, the crisis arose "unnaturally" and strongly affected the financial health of all companies. There was a danger of an immediate slump in corporate cashflows leading many firms with otherwise low credit risk to fall into insolvency and damaging the economy in the long run. To prevent job losses and a cascade of secondary insolvencies, many governments introduced measures to overcome the imminent liquidity crisis. A crucial policy question was how large the support should be. The simulation presented here gives a rough estimate of the minimum and sufficient liquidity support (Figure 7). The minimum liquidity support reflects the sum of the negative gross operating surpluses of the individual divisions of the economy. The sufficient support reflects the difference between gross operating surplus in the simulation and adjusted gross operating surplus in 2019Q4.[15]

It can be presumed that the current crisis will also affect corporate credit quality and cause a higher amount of loan defaults. The default rate forecasts were performed in two versions. Version A assumes no government support measures, while version B takes into account all the measures announced up to 15 April 2020.[16] The aggregate default rates are displayed in Figure 8.

[15] Note that these estimates are based on the aggregated values for the individual divisions. Such aggregation may cause them to be underestimated. Imagine a naive example division consisting of two firms: A and B. If firm A's operating profit is equal to +1 and firm B's is equal to -1, the aggregate operating profit of the division is 0. However, the liquidity need is driven only by the sum of the negative operating profits, which in our example is equal to -1.

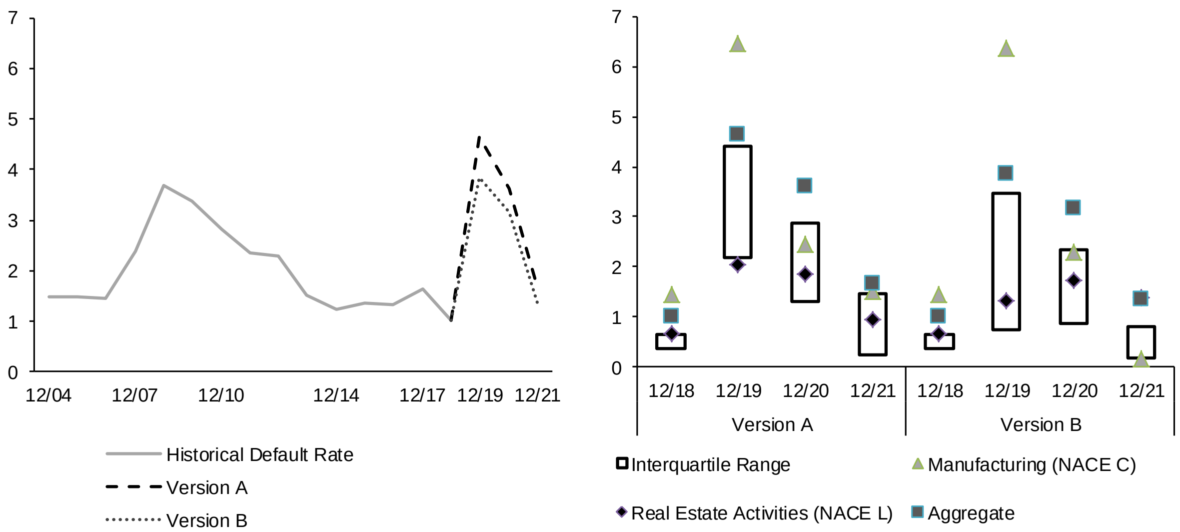
[16] On 15 April, the sum of the measures – consisting of direct financial support, postponement of tax and social contributions, and a "kurzarbeit" programme – were budgeted at about CZK 90 billion (2% of GDP). Moreover, a loan moratorium accounts for another CZK 65 billion of quarterly cashflow. However, this measure had been taken up by approximately 15% of firms as of 15 April. The support packages were applied in two modes: measures of a direct support nature were added to the gross operating surplus of the individual divisions in the targeted periods without further consequences, while deferral-type measures were first added to gross operating surplus in the designated period and deducted afterwards. Note also that the support measures affect not only the profit/cash flow of firms on a divisional level, but also scenario variables such as credit growth and property price growth (as the measures include the abrogation of property transfer tax).

Figure 7: Minimum and Sufficient Liquidity Support in Covid-19 Scenario (in % of Nominal GDP in 2019)



Source: Author's calculations

Figure 8: Default Rates Implied by Covid-19 Scenario (in %)



Source: Historical default rates – CNB; forecasts – author's calculations

The Covid-19 scenario implies a sharp spike in loan defaults in 2020, a slight decrease in 2021 and a return to low levels in 2022. The difference between version A and version B shows the economic effect of the government measures. With the application of support measures, the cumulative amount of loan defaults over the three-year horizon drops by 1.6 p.p., from 10% to 8.4%. From a more detailed viewpoint, the effect of the fiscal measures is largest in transportation and storage, construction and ICT. From a financial sector perspective, 8.4% of loan defaults will result in significant credit losses. As shown in Section 4.2, by making an assumption about the LGD parameter, regulators can use these findings well in advance to estimate the capital at risk and the sufficient level of capital. Note also that the presented defaults are based solely on economic developments. The actual occurrence of defaults may have been postponed to 2021 due to deferred loan payments/the loan moratorium, because banks do not need to treat these deferred loans as being in default, while probably some proportion of them are already in default.

5. Conclusion

This paper provides a framework for conducting simulations and stress testing in the non-financial corporate sector and shows some of the possible outcomes, such as the sectional default rate predictions implied by the simulation results and the desired liquidity support during the current crisis. The presented approach can be used to support policy decisions ranging from those which need to be made urgently, such as plugging possible liquidity gaps and determining the desired size of government loan guarantees and central bank support programmes ("funding for lending" and the like), to longer-term challenges such as concentration risk issues and the calibration of optimal capital buffers. On a more conceptual and forward-looking note, by constructing suitable scenarios and incorporating additional data about the emissions intensity of individual divisions, the presented framework can capture and quantify possible risks to the banking sector connected with the transformation to a green and low-carbon economy and the impact of that process at the level of economic divisions.

The proposed approach strongly depends on the underlying scenario, which is strictly exogenous and can be taken from fiscal, monetary or international institutions, as they perform detailed economic predictions on a regular basis. Alternatively, as shown in Section 4.3, the scenario can be created flexibly in real time at the beginning of a crisis, with the advantage of providing early yet reliable results. Another benefit of the proposed approach is the possibility to differentiate the shock intensity across the parts of the economy.

The methodology described in this paper opens up space for increasing or reducing the number of variables included in the learning process. Some of the variables used may be difficult to obtain in some countries, and others can be easily redefined or expanded. The learning process will always extract the most influential ones. In some situations, lasso regularization can be replaced with the elastic net for enhanced interpretability.

The results contain reliable historical estimates of corporate default rates for the main sections of the Czech economy in the period 2004–2018. They also offer reasonable default rate forecasts in both presented scenarios almost in real time during the crisis. Finally, other outcomes presented in the paper, including the range of additional liquidity support needed and an evaluation of the measures implemented, contribute to the policy debate.

References

AGARWAL, V. AND R. TAFFLER (2008): "Comparing the Performance of Market-based and Accounting-based Bankruptcy Prediction Models." *Journal of Banking & Finance*, 32:1541-1551.

ALANISTE CONTRERAS, M. AND G. FAGIOLO, (2014): "Propagation of Economic Shocks in Input-Output Networks." LEM Working Paper Series, 2014/09.

BAUM, C. F., M. CAGLAYAN, AND J. T. BARKOULAS (2001): "Exchange Rate Uncertainty and Firm Profitability." *Journal of Macroeconomics*, 23(4):565-576.

BLACKWOOD, G. J. (2017): "Same Shock, Separate Channels: House Prices and Firm Performance in the Great Recession." University of Maryland at College Park.

BCBS (2000): "Principles for the Management of Credit Risk." Basel Committee on Banking Supervision.

CASTRÉN, O., S. DÉES, AND F. ZAHER (2009): "Stress-testing Euro Area Corporate Default Probabilities Using a Global Macroeconomic Model." *Journal of Financial Stability*, 6(2):64-78.

CIPOLLINI, A. AND G. MISSAGLIA (2008): "Forecasting Industry Sector Default Rates Through Dynamic Factor Models." *Journal of Risk Model Validation*, 2(3):59-79.

DIXON, P. AND D. W. JORGENSON (2013): "Handbook of Computable General Equilibrium Modeling." 1A and 1B.

DREHMANN, M. (2005): "A Market Based Macro Stress Test for the Corporate Credit Exposures of UK Banks." Bank of England, mimeo.

EUROSTAT (2006): "NACE Rev. 2 - Statistical Classification of Economic Activities." Methodologies and working papers.

FEYEN, E., N. FIESS, I. ZUCCARDI HUERTAS, AND L. PILLONCA (2017): "Which Emerging Markets and Developing Economies Face Corporate Balance Sheet Vulnerabilities?" Policy Research Working Paper Series, No. 8198, World Bank.

FIGLEWSKI, S., H. FRYDMAN, AND W. LIANG (2012): "Modeling the Effect of Macroeconomic Factors on Corporate Default and Credit Rating Transitions." *International Review of Economics and Finance*, 21(1):87-105.

FLOM, P. (2018): "Stopping Stepwise: Why Stepwise Selection Is Bad and What You Should Use Instead." Available at: https://towardsdatascience.com/stopping-stepwise-why-stepwise-selection-is-bad-and-what-you-should-use-instead-90818b3f52df [4 April 2020].

FRAIT, J. AND Z. KOMARKOVA (2012): "Macroprudential Policy and Its Instruments in a Small EU Economy." CNB Research and Policy Notes, 3/2012.

FRIEDMAN, J., T. HASTIE, AND R. TIBSHIRANI (2016): "Regularization Paths for Generalized Linear Models via Coordinate Descent." *Journal of Statistical Software*, 33(1): 1-22.

HARRELL, F. (2001): "Regression Modeling Strategies with Applications to Linear Models, Logistic Regression and Survival Analysis." New York: Springer-Verlag.

HOHN, K. (2004): "Application to Economics: Leontief Model." Kansas State University.

KARASULU, M. AND M. T. JONES (2006): "The Korean Crisis: What Did We Know and When Did We Know It? What Stress Tests of the Corporate Sector Reveal." IMF Working Papers, 06/114.

KLEIN, N. (2016): "Corporate Sector Vulnerabilities in Ireland." IMF Working Papers, 16/211.

LEE, K.-M., J.-S. YANG, G. KIM, J. LEE, K.-I. GOH, AND I.-M. KIM (2011): "Impact of the Topology of Global Macroeconomic Network on the Spreading of Economic Crises." PLoS ONE, 6(3):

LEONTIEF, W. (1936): "Quantitative Input and Output Relations in the Economic System of the United States." *Review of Economics and Statistics*, 18:105-125.

LEQUILLER, F. AND D. BLADES (2014): "Understanding National Accounts: Second Edition." OECD Publishing.

LOUZIS, D. P., A. T. VOULDIS, AND V. L. METAXAS (2012): "Macroeconomic and Bank-specific Determinants of Non-performing Loans in Greece: A Comparative Study of Mortgage, Business and Consumer Loan Portfolios." *Journal of Banking & Finance,* 36:1012-1027.

PANOŠ, J. AND P. POLÁK (2020): "How to Improve the Model Selection Procedure in a Stress-testing Framework." CNB Working Papers, 9/2019.

RADIVOJEVIC, N. AND J. JOVOVIC (2017): "Examining of Determinants of Non-Performing Loans." *Prague Economic Papers*, 2017(3):300-316.

SIMONS, D. AND F. ROLWES (2009): "Macroeconomic Default Modelling and Stress Testing." *International Journal of Central Banking*, 5(3):177-204.

TIBSHIRANI, R. (1996): "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society*, 58:267-288.

VIROLAINEN, K. (2004): "Macro Stress Testing with a Macroeconomic Credit Risk Model for Finland." Bank of Finland Discussion Papers, 18/2004.

Appendix

Figure A1: Production and Gross Operating Surplus of Selected Sections and Divisions in Deep Recession Scenario (Real Terms Index: 2019Q2 = 1)

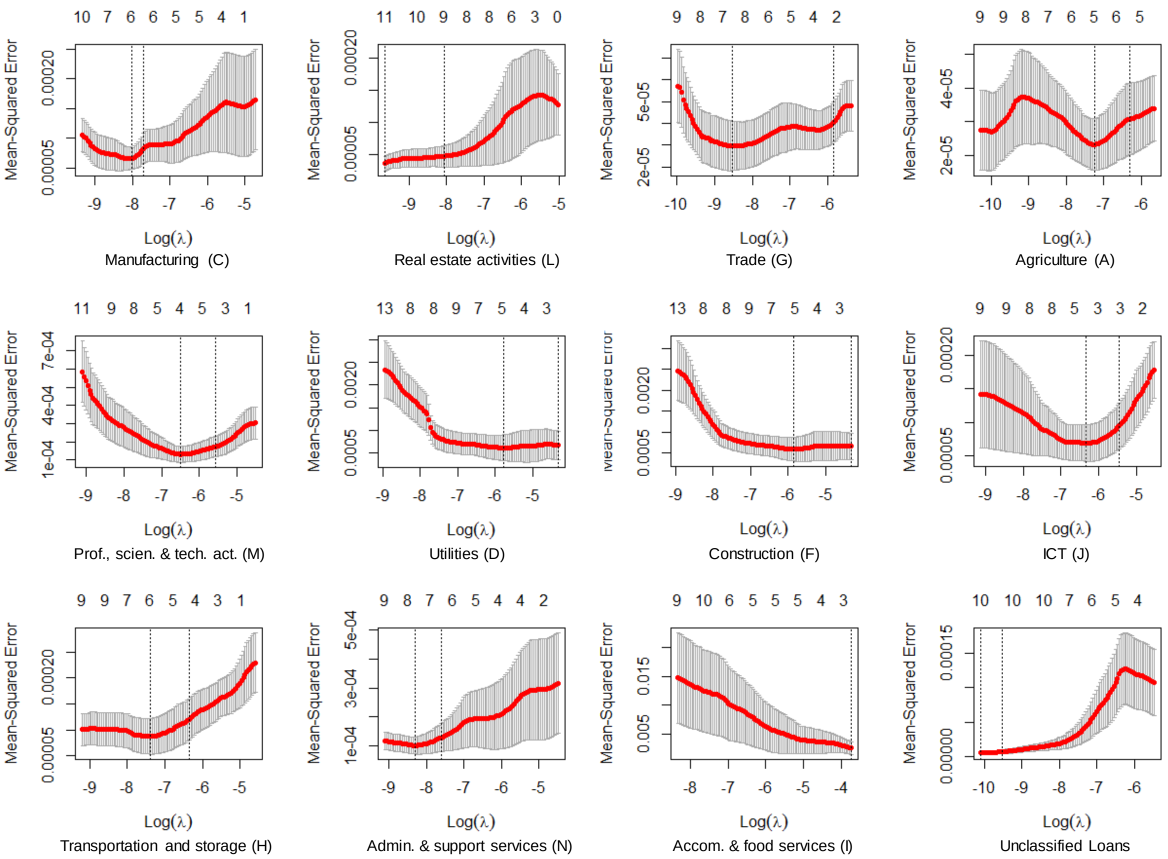
Note: A – Agriculture, forestry and fishing, C – Manufacturing, D – Electricity, gas, steam and air conditioning supply, sewerage and waste management, F – Construction, G – Wholesale and retail trade, repair of motor vehicles and motorcycles, H – Transportation and storage, J – Information and communication, L – Real estate activities, M – Professional, scientific and technical activities, N – Administrative and support service activities. For illustration purposes, the chart also displays divisions classified under C – Manufacturing industry: NACE 20 – Manufacturing of chemicals and chemical products, NACE 22 – Manufacture of rubber and plastic products and NACE 29 – Manufacture of motor vehicles, trailers and semi-trailers.

Figure A2: Algorithm Adjustment for Capturing Supply-side Shock (from R)

```
if (d_S(t) > d_Sf(t))
    {d_S(t) <- d_Sf(t)
     Intermediate_Cons(t) <- Intermediate_Cons(t-1) + (d_S(t))*A
     Final_demand(t) <- Final_demand(t-1) + (d_S(t))*(1-rowSums(A))
    }
```

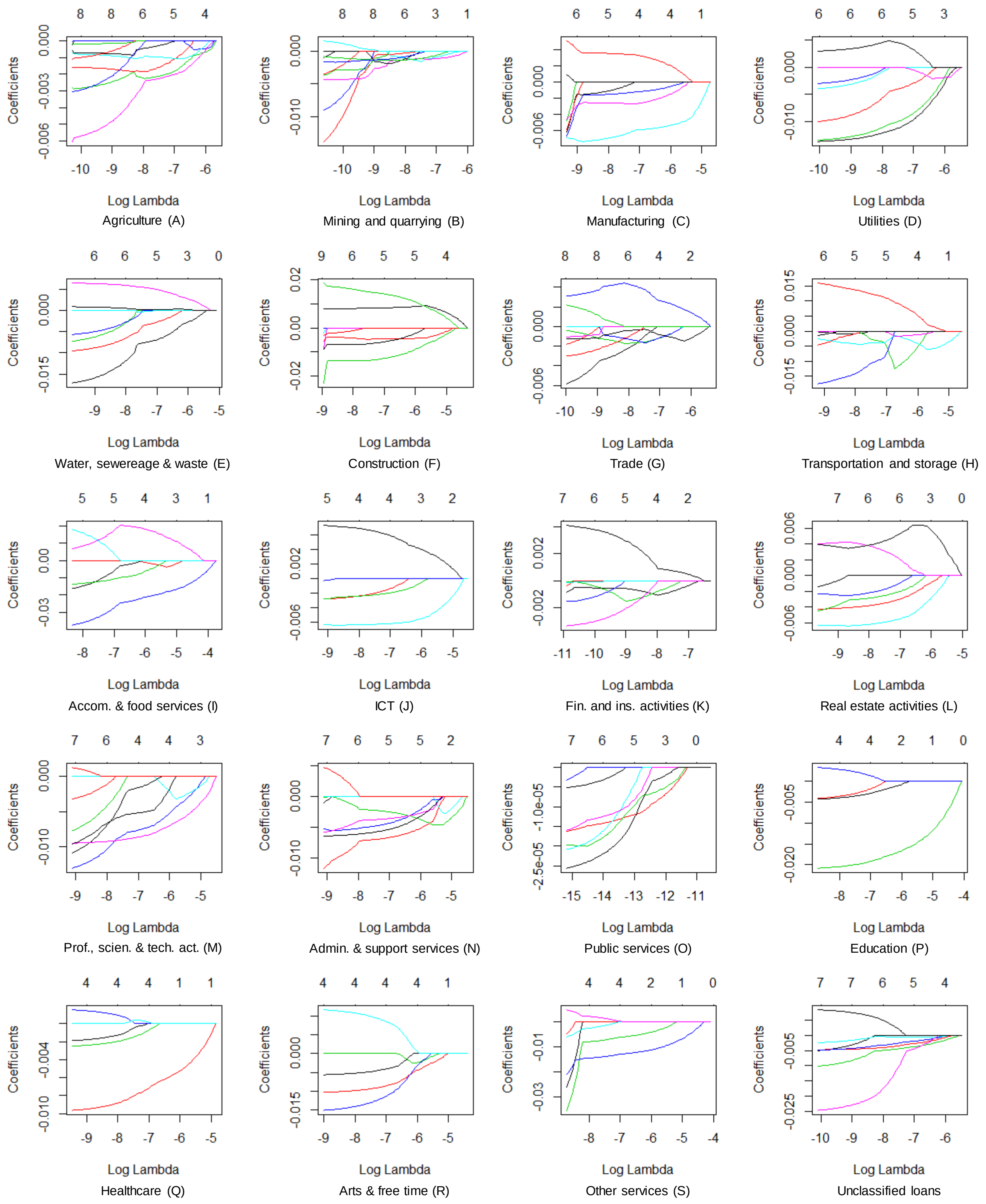
Note: d_S(t) stands for change of supply implied by demand shock, d_Sf(t) represents primary assumption about supply shock and A is technology matrix (defined on page 3).

Figure A3: Cross-validation Characteristics for Major Sections



Note: Figure shows mean squared error and its distribution during cross-validation dependent on value of lambda and number of chosen parameters.

Figure A4: Lambda Traceplots for All Sections



Note: Figure shows numbers of non-zero coefficients and their values according to value of lambda.

CNB Working Paper Series (since 2019)

WP 12/2020	Vojtěch Siuda	*A top-down stress-testing framework for the nonfinancial corporate sector*
WP 11/2020	Hana Hejlová Michal Hlaváček Blanka Vačkova	*Estimating commercial property price misalignment in the CEE countries*
WP 10/2020	Zuzana Rakovská Dominika Ehrenbergerová Martin Hodula	*The power of sentiment: Irrational beliefs of households and consumer loan dynamics*
WP 9/2020	Ivan Sutóris	*The intertemporal cost of living and dynamic inflation: The case of the Czech Republic*
WP 8/2020	Martin Hodula Jan Janků Martin Časta Adam Kučera	*On the determinants of life and non-life insurance premiums*
WP 7/2020	František Brázdik Tibor Hlédik Zuzana Humplová Iva Martonosi Karel Musil Jakub Ryšánek Tomáš Šestořád Jaromír Tonner Stanislav Tvrz Jan Žáček	*The g3+ model: An upgrade of the Czech National Bank's core forecasting framework*
WP 6/2020	Volha Audzei Jan Brůha	*A model of the Euro Area, China and the United States: Trade links and trade wars*
WP 5/2020	Dominika Ehrenbergerová Martin Hodula Zuzana Rakovská	*Does capital-based regulation affect bank pricing policy?*
WP 4/2020	Alexis Derviz	*Sovereign capital, external balance, and the investment-based Balassa-Samuelson effect in a global dynamic equilibrium*
WP 3/2020	Milan Szabo	*Growth-at-risk: Bayesian approach*
WP 2/2020	Martin Hodula Ngoc Anh Ngo	*Finance, growth and (macro)prudential policy: European evidence*
WP 1/2020	Michal Franta Ivan Sutóris	*Dynamics of Czech inflation: The role of the trend and the cycle*
WP 12/2019	Josef Bajzík Tomáš Havránek Zuzana Iršová Jiří Schwarz	*The elasticity of substitution between domestic and foreign goods: A quantitative survey*

WP 11/2019	Martin Hodula Simona Malovaná Jan Frait	*Too much of a good thing? Households' macroeconomic conditions and credit dynamics*
WP 10/2019	Martin Hodula Simona Malovaná Jan Frait	*Introducing a new index of households' macroeconomic conditions*
WP 9/2019	Jiří Panoš Petr Polák	*How to improve the model selection procedure within a stress testing framework?*
WP 8/2019	Sebastian Gechert Tomáš Havránek Zuzana Iršová Dominika Kolcunová	*Death to the Cobb-Douglas production function? A quantitative survey of the capital-labor substitution elasticity*
WP 7/2019	Alexis Derviz	*Coexistence of physical and crypto assets in a stochastic endogenous growth model*
WP 6/2019	Dominika Ehrenbergerová Simona Malovaná	*Introducing macro-financial variables into semi-structural model*
WP 5/2019	Martin Hodula	*Monetary policy and shadow banking: Trapped between a rock and a hard place*
WP 4/2019	Simona Malovaná Žaneta Tesařová	*Banks' credit losses and provisioning over the business cycle: Implications for IFRS 9*
WP 3/2019	Aleš Bulíř Jan Vlček	*Monetary policy is not always systematic and data-driven: Evidence from the yield curve*
WP 2/2019	Dominika Kolcunová Simona Malovaná	*The effect of higher capital requirements on bank lending: The capital surplus matters*
WP 1/2019	Jaromír Baxa Tomáš Šestořád	*The Czech exchange rate floor: Depreciation without inflation?*

CNB Research and Policy Notes (since 2019)

RPN 3/2020	Simona Malovaná Martin Hodula Zuzana Rakovská	*Researching the research: A central banking edition*
RPN 2/2020	Simona Malovaná Josef Bajzík Dominika Ehrenbergerová Jan Janků	*A prolonged period of low interest rates: Unintended consequences*
RPN 1/2020	Simona Malovaná	*How to organize research in central banks: The Czech National Bank's experience*

RPN 4/2019	Tomáš Konečný Lukáš Pfeifer	*Macroprudential ring-fencing*
RPN 3/2019	Petr Polák Jiří Panoš	*Impact of expectations on IFRS 9 loan loss provisions*
RPN 2/2019	Jan Filáček Ivan Sutóris	*Inflation targeting flexibility: The CNB's reaction function under scrutiny*
RPN 1/2019	Iveta Polášková Luboš Komárek Michal Škoda	*The contemporary role of gold in central banks' balance sheets*

CZECH NATIONAL BANK
Na Příkopě 28
115 03 Praha 1
Czech Republic

ECONOMIC RESEARCH DIVISION
Tel.: +420 224 412 321
Fax: +420 224 412 329
http://www.cnb.cz
e-mail: research@cnb.cz

ISSN 1803-7070