

# STRESS TESTING OF NON-FINANCIAL CORPORATE SECTOR: A TOP-DOWN INPUT-OUTPUT FRAMEWORK\*

Vojtěch Siuda <sup>a</sup>

## Abstract

This paper provides a framework for conducting simulations and stress testing in the non-financial corporate 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 entirely 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 forecasting portfolio default rates – are shown in the paper. The historical default rate estimates obtained are accurate and economically sensible for most industries and exhibit high reliability even under 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.

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

**JEL Classification:** G01, G32, H63

## Introduction

Given its share in gross value added, the non-financial corporate sector forms the backbone of all economies. However, firms often operate with significant leverage and

---

a Czech National Bank, Financial Stability Department, and Prague University of Economics and Business, Faculty of Finance and Accounting, Prague, Czech Republic  
E-mail: vojtech.siuda@cnb.cz

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 20<sup>th</sup> and the first two decades of the 21<sup>st</sup> century. The global economy experienced a strong negative demand shock during the Great Recession in the late 2000s. This spilt over 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 inter- and intra-sectoral links and assessing 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 as follows. 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 identifying a varying set of determinants of sectional default rates, where the determinants include industry-specific indicators of economic performance and profitability (Section 4.1). Combining 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 is organized as follows. First, some relevant literature on modelling aggregate credit losses and its main findings are provided. The next section describes the data and the stress-testing methodology, consisting of economic simulation using input-output tables and a machine learning algorithm for default rate modelling.

The following 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 the first level – letters A–U – and these are further divided into 88 *divisions* on the second level – two-digit code). Some additional outcomes that can serve as inputs for policy discussion are also presented.

## 1. Related Literature

Stress testing of the banking sector's solvency became a standard regulatory tool in the late 1990s<sup>1</sup>. It developed apace after the global financial crisis, when stress tests also began to assess 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 crucial feature of shock propagation. The macroeconomic determinants of the corporate default rate on the aggregate level were analysed by Karasulu and Jones (2006). They 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 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 macro-financial 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

---

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.

models and showed robust results under different economic conditions. However, this approach seems difficult for stress-testing purposes and 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 Merton structural credit 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 to 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.

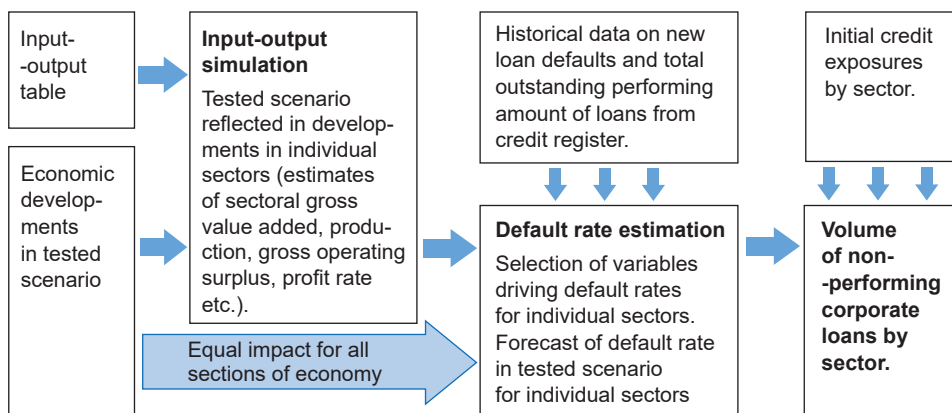
Several studies strive to explain and estimate the effects of loan defaults on banks' balance sheets. The determinants of loan loss provisions and non-performing loan ratios were analysed, for example, by Louzis *et al.* (2012), Radivojević and Jovović (2017) and Frait and Komárková (2012). Contrary 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.

As for corporate credit risk estimation in the Czech Republic, Kadlčáková and Keplinger (2004) created a credit rating system based on a database of individual Czech firms. They achieved a competitive comparison with widely used CreditMetric and CreditRisk+ approaches using industry-specific variables (*e.g.*, NACE specific credit spread). A latent-factor credit risk model for the Czech economy dependent on macroeconomic variables was performed for the NFC and household sectors by Jakubík (2007) and Jakubík and Schmieder (2008). The model based on macroeconomic variables showed a very accurate fit for the corporate default rate according to the presented findings. However, calibration on the data with a transformation crisis may lead to the high sensitivity of the default rate in the stress test. A coherent framework for stress-testing of the individual risk areas in the banking sector in Czech conditions was introduced by Geršl *et al.* (2013). A more technically oriented contribution to the credit risk in the Czech Republic was made by Panoš and Polák (2020). They did extensive work on optimal model selection in scenario-conditional default rate predictions. Our present paper follows top-down corporate credit risk estimation approaches, emphasising the default rate of individual industries (sections) based on their macroeconomic performance. This paper thus should be viewed as a complementary work that extends top-down modelling of corporate credit risk in the Czech Republic.

## 2. 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 of the economy. These results are then grouped into section-specific performance variables and selected aggregate variables and form the basis for the second step, which consists of a default rate learning process. Figure 1 shows the simplified logic of the process.

**Figure 1: Basic logic of NFCs stress-testing framework**



Note: Grey boxes represent input data, while white boxes represent calculations and value-added of the author.

Source: Autor's own elaboration

### 2.1 Input-output simulation

Input-output analysis was chosen as the primary vehicle for modelling shock transmission between the divisions of the economy. This approach is preferable to computable general equilibrium models (see, *e.g.*, Dixon and Jorgenson, 2013). The goal is to describe the shock propagation between and interactions of economic divisions rather than predict economic developments (defined in the scenario). One of the drawbacks of input-output analysis is its data-demanding nature 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 regularly presented predictions of 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 more comprehensive range of analyses of possible economic developments that are hard to capture by the current model families. However, this approach also bears a higher risk of possible errors and inherent inconsistencies of variables and requires a deep knowledge of economic theory.

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 made or imported by enterprises are consumed, stored in inventories or exported. In other words, there exists an identity where production at purchasers' prices and imports (collectively as supply) equals intermediate consumption and final demand (collectively as demand), which corresponds with the information in the production account.

To solve the open<sup>2</sup> model problem, let us denote:

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

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

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

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

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

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

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

2 One can distinguish between two models depending on the ultimate goal of the analysis. In an open model, there exists a consumption vector  $\mathbf{D} = (d_1, d_2, \dots, d_S)^T$ , 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).

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

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

The main criticism of the Leontief model relates to its static nature. 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 according to structural changes in the economy and technological progress. Moreover, they can be expected to change more significantly during stress periods. To overcome this limitation, we borrow the approach of Alatryste 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, the top-down corporate credit risk estimation approaches the technology matrix  $A'$ , where each element in the  $k$ -th row and  $k$ -th column has been updated by  $q_k$  to:

$$a'_{k,j} = q_k a_{k,j} \quad (3)$$

$$a'_{i,k} = q_k a_{i,k} \quad (4)$$

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

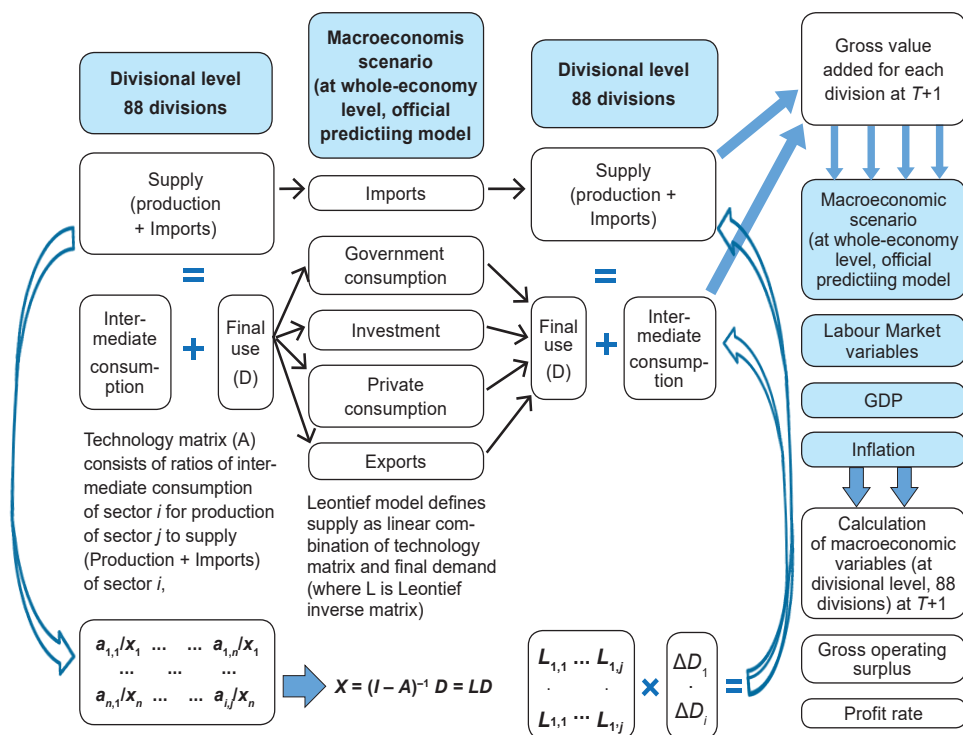
$$\Delta X = (I - A')^{-1} \Delta D = L' \Delta D. \quad (5)$$

This mechanism assumes accurate foresight of firms about economic activity. It is a self-fulfilling process that adjusts coefficients of the technology matrix, which adapts immediately and wholly to the new level of final demand.

The economic shock is entirely contained in the whole-economy scenario through the development of GDP components. 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.<sup>3</sup> Under the differentiated approach, the shock impacts on individual divisions

3 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. Still, this assumption is very strong and the uniform impact is not likely to occur in reality.

**Figure 2: Simplified one step of macrosimulation**



Note: The figure shows one step of the macrosimulation. In the begging (left side of the figure), there are known values of production, import, intermediate consumption and final use for all 88 divisions of the economy, which are obtained from the input-output table. At the same time, there is a known scenario development of economy (GDP components, labour market variables, price indices –represented in the figure generally as inflation) defined on the whole-economy level. With assumptions about the final use distribution in 88 divisions in the following period (the period is quarterly in the presented model), the Leontief model can calculate corresponding intermediate consumption and production for all divisions. Afterwards, other macroeconomic indicators can be derived as gross value-added, gross operating surplus, etc., for all 88 divisions.

Source: Autor's own elaboration

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.<sup>4</sup> This opens up the possibility of increasing the stress on some segments and reducing that on others.

With the known level of change in individual components of final demand and imports, the Leontief approach then finds a corresponding change in production and intermediate consumption for all economy divisions and that the identity holds. The simplified macroeconomic simulation process is shown in Figure 2.

The speed of adaptation of intersectoral relations has risen strongly over the past few decades. 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.

## 2.2 Default rate estimation

Based on the 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. A bank uses accounting information and a set of accounting ratios to derive a firm's credit quality and related credit risk. In our approach, the variables entering the default rate model draw on similar information converted to 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 the manufacturing industry (automotive, machinery). Still, a broader view was necessary to balance systematic patterns and pure noise-fitting.

---

4 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).

**Table 1: Accounting variables and their macroeconomic equivalents**

Macro variables	Accounting variables
Net operating surplus / Productiont	Corresponds to profits after depreciation and taxes normalized by sales (Net earnings / Sales)
Gross operating surplus / Gross value addedt	(EBITDA / Gross profit)
Gross operating surplus / Productiont	Corresponds to gross operating profit to sales (EBITDA/Sales)
Gross operating surplus / Average interest rate × volume of credit	Close to interest coverage ratio (EBITDA / Interest expenses)
Gross value addedt / Gross value addedt-1	Corresponds to gross profit dynamics

Source: Autor's own elaboration

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 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 non-financial corporations (Blackwood (2018) describes the link between property prices and firm performance; for the exchange rate, see, *e.g.*, Baum *et al.*, 2001).<sup>5</sup>

---

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

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.<sup>6</sup> Table 2 summarizes the set of all the variables initially entering the default rate modelling, including the logical sign restrictions imposed on the coefficient 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 as 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 primary 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 dichotomy between the standard credit risk model and the model presented here lies in the very nature of these approaches. The present 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 financial sector's stability under severe conditions (tail events) and to help answer crucial policy questions. A model with a 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 the parameter values. Predictions of the aggregate amount of loan defaults in each section of the economy in the given scenario are then the outcome of the model.

There are several basic approaches to variable selection. Notably, many more complex approaches exist for optimal model and variable selection in default rate forecasting. 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, the stepwise selection remains the most common credit risk method despite all the related issues (as summarized, *e.g.*, by Harrell, 2001 or Flom, 2018). Nevertheless, instead of stepwise methods, we employed the LASSO estimator (Tibshirani, 1996) to select the optimal set of predictors and regularize the problem. This is a machine learning algorithm that shrinks the parameters of less significant variables to 0.

---

6 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.

**Table 2: List of all section-specific and aggregate variables initially entering default rate learning algorithm**

	Variable	Abbreviation	Restriction
<b>Dependent variable</b>	Default rate <sub>t</sub> => default rate at time t	DF <sub>t</sub>	
<b>Section-specific explanatory variables</b>	Constant	Const.	
	Default rate <sub>t-1</sub> => default rate at time t-1 (AR1 process)	DF <sub>t-1</sub>	>0
	Net operating surplus <sub>t</sub> / Production <sub>t</sub>	(NOS/P) <sub>t</sub>	<0
	Net operating surplus <sub>t-1</sub> / Production <sub>t-1</sub>	(NOS/P) <sub>t-1</sub>	<0
	Gross operating surplus <sub>t</sub> / Gross value added <sub>t</sub>	(GOS/GVA) <sub>t</sub>	<0
	Gross operating surplus <sub>t-1</sub> / Gross value added <sub>t-1</sub>	(GOS/GVA) <sub>t-1</sub>	<0
	Gross operating surplus <sub>t</sub> / Production <sub>t</sub>	(GOS/P) <sub>t</sub>	<0
	Gross operating surplus <sub>t-1</sub> / Production <sub>t-1</sub>	(GOS/P) <sub>t-1</sub>	<0
	(Net operating surplus <sub>t</sub> – Net operating surplus <sub>t-1</sub> ) / Production <sub>t-1</sub>	Δ(NOS/P) <sub>t</sub>	<0
	(Net operating surplus <sub>t-1</sub> – Net operating surplus <sub>t-2</sub> ) / Production <sub>t-2</sub>	Δ(NOS/P) <sub>t-1</sub>	<0
	(Gross operating surplus <sub>t</sub> – Gross operating surplus <sub>t-1</sub> ) / Gross value added <sub>t-1</sub>	Δ(GOS/GVA) <sub>t</sub>	<0
	(Gross operating surplus <sub>t-1</sub> – Gross operating surplus <sub>t-2</sub> ) / Gross value added <sub>t-2</sub>	Δ(GOS/GVA) <sub>t-1</sub>	<0
	(Gross operating surplus <sub>t</sub> – Net operating surplus <sub>t-1</sub> ) / Production <sub>t-1</sub>	Δ(GOS/P) <sub>t</sub>	<0
	(Gross operating surplus <sub>t-1</sub> – Net operating surplus <sub>t-2</sub> ) / Production <sub>t-2</sub>	Δ(GOS/P) <sub>t-1</sub>	<0
	Gross value added <sub>t</sub> / Gross value added <sub>t-1</sub>	ΔGVA <sub>t</sub>	<0
	Gross value added <sub>t-1</sub> / Gross value added <sub>t-2</sub>	ΔGVA <sub>t-1</sub>	<0
	Gross operating surplus <sub>t</sub> / Interest paid <sub>t</sub>	ICR <sub>t</sub>	
	Gross operating surplus <sub>t-1</sub> / Interest paid <sub>t-1</sub>	ICR <sub>t-1</sub>	
<b>Aggregate explanatory variables</b>	Property prices <sub>t</sub> / Property prices <sub>t-1</sub>	ΔPP <sub>t</sub>	
	Property prices <sub>t-1</sub> / Property prices <sub>t-2</sub>	ΔPP <sub>t-1</sub>	
	FX <sub>t</sub> / FX <sub>t-1</sub> *	ΔFX <sub>t</sub>	
	*Nominal FX rate is in direct quotation => higher value means relative depreciation		

Source: Autor's own elaboration

Consider a vector of observations  $Y = (y_1, y_2, \dots, y_T)$  of length  $T$  and a  $\mathbf{T} \times \mathbf{P}$  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:

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

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

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

where  $\lambda$  is the tuning parameter that regulates the degree of shrinkage.<sup>7</sup> Generally, the superiority of regularization techniques over a 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 present approach, where the defined explanatory variables tend to be correlated and shrink the less significant ones. The preference for LASSO penalization over ridge regression<sup>8</sup> 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 likely to be 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 should not be a source of growth in default rates) to ensure the robustness of estimates and improve the stability of parameters. In the case of interest coverage ratio (ICR), conventional logic concludes that with rising ICR, the default rate should decrease (resulting in a negative ICR coefficient). However, during a crisis period, interest rates usually decrease, which could outweigh the slump in profits, and the final ICR can be higher despite higher default occurrence. Due to this effect, there were no logical restrictions for the ICR variable. The algorithm was performed in R with the glmnet package (Hastie and Qian, 2016). The optimal value of the tuning parameter  $\lambda$  for each section was optimized by cross-validation.<sup>9</sup>

7 Obviously, lambda equalling zero returns a standard linear regression.

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

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

## 2.3 Data

Input-output tables (available from the Czech Statistical Office) were used for the simulation. The simulation itself starts from the last known input-output table<sup>10</sup> (for 2018 at the time of writing). This table contains data on the components of gross value added (costs of labour, net taxes, consumption of fixed capital, operating surplus and mixed income) by each division of the economy, inter-divisional relations (intermediate consumption) and final demand. Moreover, time series of these variables from 2002 to 2018 were created from historical input-output tables. These time series were used for the learning process in modelling the historical default rates as described in Sections 3.2 and 4.1. Real estate prices (available from the Czech Statistical Office<sup>11</sup>), EUR/CZK exchange rate<sup>12</sup> and corporate loan interest rate data<sup>13</sup> (available from the Czech National Bank) were also used for modelling the historical default rates.

Historical default rate data are based on flows of new non-performing loans and total volumes of performing loans. These were taken from the Czech Republic's Central Credit Register managed by the Czech National Bank.<sup>14, 15</sup> These data were used to calculate historical default rates for each section of the economy. Default rates are defined in a forward-looking manner as:

$$\text{Default rate}_{t,i} = \frac{NDL_{t+1,i}}{TOPA_{t,i}} \quad (8)$$

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

For demonstration of results, we chose two economic scenarios: the Deep Recession scenario and the COVID-19 scenario. The Deep Recession scenario corresponds

---

10 Historical and current input-output tables of the Czech economy are available at [https://apl.czso.cz/pil/rocenka/rocenkaout.dod\\_uziti?mylang=EN](https://apl.czso.cz/pil/rocenka/rocenkaout.dod_uziti?mylang=EN)

11 House price index: <https://www.czso.cz/csu/czso/house-price-index-2-quarter-of-2021>

12 Available at <https://www.cnb.cz/en/financial-markets/foreign-exchange-market/central-bank-exchange-rate-fixing/central-bank-exchange-rate-fixing/>

13 Available at [https://www.cnb.cz/cnb/STAT.ARADY\\_PKG.VYSTUP?p\\_period=1&p\\_sort=2&p\\_des=50&p\\_sesuid=58863&p\\_uka=17&p\\_strid=AAABBA&p\\_od=200101&p\\_do=202108&p\\_lang=EN&p\\_format=0&p\\_decsep](https://www.cnb.cz/cnb/STAT.ARADY_PKG.VYSTUP?p_period=1&p_sort=2&p_des=50&p_sesuid=58863&p_uka=17&p_strid=AAABBA&p_od=200101&p_do=202108&p_lang=EN&p_format=0&p_decsep)

14 These data are not available to general public. The Central Credit Register contains loan data of non-financial corporations on an individual basis.

15 Due to the data quality and sufficient granularity, only bank loans are considered even if they account only for around half of the total NFCs debt. There is no database for Czech in-house funding, financial leasing or cross-border financing, which also includes defaulted loans and the data do not provide a NACE code or another type of industry identification. According to bond financing, there were no defaulted bonds traded on regulated markets in recent years.

to the Adverse Scenario published in *Risks to Financial Stability and Their Indicators 2019* by the Czech National Bank. This was used to simulate financial developments and financial sector resiliency in a period of a strong economic recession. The variables used for the input-output simulation and obtaining section-specific variables are mutually consistent, model-based and include:<sup>16</sup>

- a) Quarterly data on private consumption, investments, government consumption, exports and imports and their price indices.
- b) Quarterly labour market variables, including nominal wage growth and unemployment rate.

Besides the section-specific variables, the following macroeconomic variables were also used for default rates predictions for Deep Recession:

- c) Quarterly corporate credit growth, interest rates and property prices.
- d) Banks' total corporate credit for the individual divisions, including its maturity structure and interest rates.

### 3. Results for the Czech Republic

#### 3.1 Historical default rate estimates

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

Interest payments (12 non-zero coefficients at the time  $t$  and 11 at the time  $t-1$ ) were selected most frequently for the individual sections. However, the mixed coefficients for ICR confirm the hypothesis that its impact varies for individual industries. The second most often chosen variable was 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 quickly when profits deteriorate (estimated coefficients are shown in Table 3).

---

16 Some of these variables may be found here: [https://www.cnb.cz/export/sites/cnb/cs/financni-stabilita/.galleries/rizika\\_pro\\_fs/rizika\\_pro\\_financni\\_stabilitu\\_a\\_jejich\\_indikatory\\_prosinec\\_2019.xlsx](https://www.cnb.cz/export/sites/cnb/cs/financni-stabilita/.galleries/rizika_pro_fs/rizika_pro_financni_stabilitu_a_jejich_indikatory_prosinec_2019.xlsx). However, some of these variables are not available to general public.

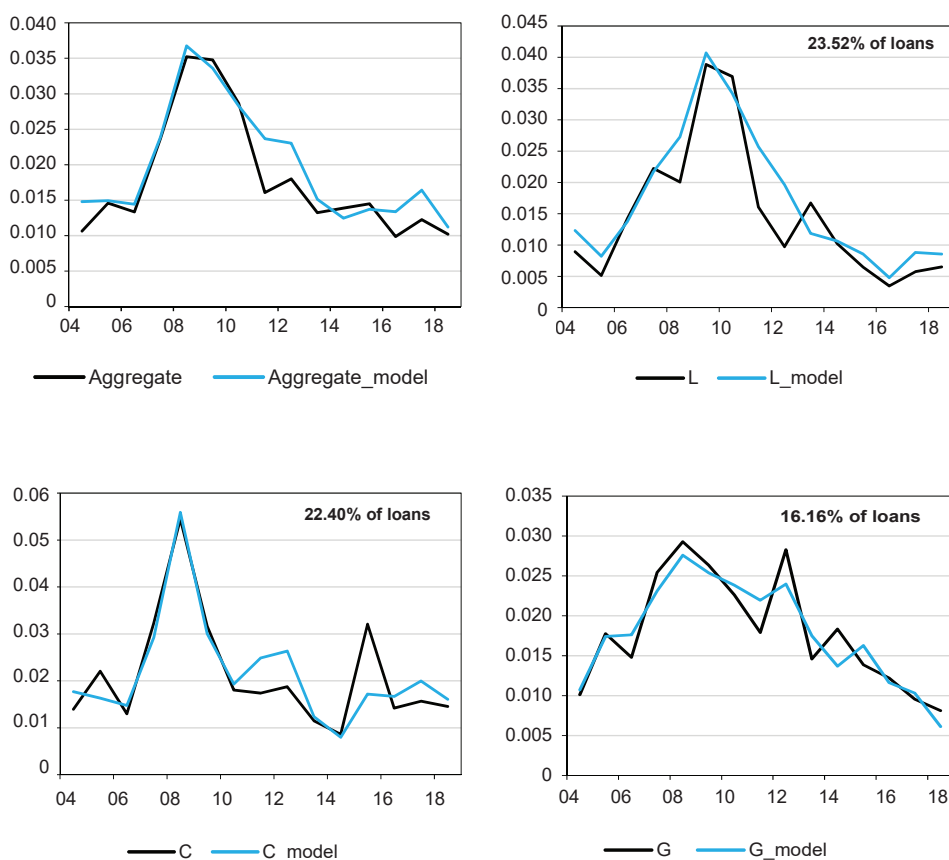
**Table 3: Coefficients obtained for individual sections**

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 <sub>t-1</sub>	0	0	0	0	0	0.009	0	0	0	0.007
$\Delta(\text{NOS}/P)_t$	-0.001	-0.001	-0.002	-0.012	0	-0.004	0	0	0	-0.001
$\Delta(\text{NOS}/P)_{t-1}$	-0.002	-0.014	0	-0.006	0	-0.013	-0.003	0	0	-0.002
$\Delta(\text{GOP}/\text{GVA})_t$	0	0	0	-0.011	0	0	0	0	-0.001	0
$\Delta(\text{GOP}/\text{GVA})_{t-1}$	0	0	-0.002	-0.001	0	0	0	0	0	0
$\Delta(\text{GOP}/P)_t$	0	-0.004	0	0	0	0	-0.001	-0.005	0	0
$\Delta(\text{GOP}/P)_{t-1}$	0	-0.009	0	-0.001	0	0	0	0	0	0
$(\text{NOS}/P)_t$	0	0	0	0	0	0	-0.001	0	0	0
$(\text{NOS}/P)_{t-1}$	-0.003	-0.004	0	0	0	0	0	-0.018	0	0
$(\text{GOP}/\text{GVA})_t$	0	0	0	0	0	0	0	0	0	0
$(\text{GOP}/\text{GVA})_{t-1}$	-0.001	0	0	0	0	0	0	0	0	0
$(\text{GOP}/P)_t$	-0.006	0	0	0	0	0	0	0	0	0
$(\text{GOP}/P)_{t-1}$	0	0	0	0	0	0	0	0	0	0
$\Delta\text{GVA}_t$	-0.001	0	-0.007	0	0	0	-0.001	-0.003	0	0
$\Delta\text{GVA}_{t-1}$	0	-0.004	-0.003	0	0	-0.006	-0.002	0	-0.016	0
$\Delta\text{PP}_t$	0	0	0	0	0	0	0	0	0	0
$\Delta\text{PP}_{t-1}$	0	0	0	0	0	-0.004	0	0	0	0
$\Delta\text{FX}_t$	0	-0.001	0	0	0	0	0	0	0	0
ICR <sub>t</sub>	-0.003	-0.002	0	0.004	0	0	0.001	-0.001	0	-0.006
ICR <sub>t-1</sub>	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 <sub>t-1</sub>	0	0	0	0	0	0	0	0	0	0
$\Delta(\text{NOS}/P)_t$	0	0	-0.002	0	0	0	0	0	-0.025	0
$\Delta(\text{NOS}/P)_{t-1}$	0	-0.001	0	0	0	0	-0.007	0	-0.006	-0.005
$\Delta(\text{GOP}/\text{GVA})_t$	0	0	0	-0.011	0	0	0	0	0	0
$\Delta(\text{GOP}/\text{GVA})_{t-1}$	0	0	-0.008	0	0	0	0	0	0	0
$\Delta(\text{GOP}/P)_t$	0	-0.004	0	0	0	0	0	0	-0.034	-0.005
$\Delta(\text{GOP}/P)_{t-1}$	0	-0.004	0	-0.005	0	0	0	0	-0.022	-0.01
$(\text{NOS}/P)_t$	0	0	0	0	0	0	0	0	0	0
$(\text{NOS}/P)_{t-1}$	0	0	0	0	0	0	0	0	0	0
$(\text{GOP}/\text{GVA})_t$	0	0	0	0	0	0	0	0	0	0
$(\text{GOP}/\text{GVA})_{t-1}$	0	0	0	0	0	0	0	0	0	0
$(\text{GOP}/P)_t$	0	0	0	0	0	0	0	0	0	0
$(\text{GOP}/P)_{t-1}$	0	0	0	0	0	0	0	0	0	0
$\Delta\text{GVA}_t$	0	0	-0.009	0	0	-0.006	0	-0.005	0	-0.005
$\Delta\text{GVA}_{t-1}$	0	0	0	-0.006	0	0	-0.001	-0.002	0	-0.002
$\Delta\text{PP}_t$	0	-0.002	0	0	0	0	0	0	0	0
$\Delta\text{PP}_{t-1}$	0	-0.006	0	0	0	0	0	0	0	0
$\Delta\text{FX}_t$	0	0	0	0	0	0	0	0	0	0
ICR <sub>t</sub>	0	0.004	-0.005	-0.006	0	0	0	-0.004	-0.006	-0.025
ICR <sub>t-1</sub>	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

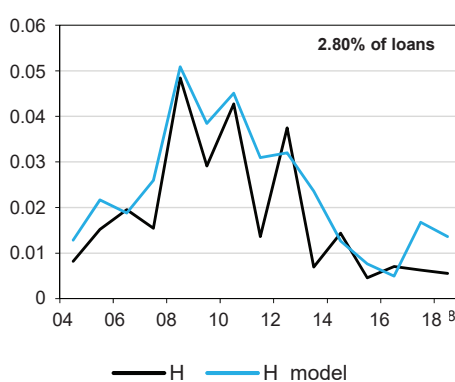
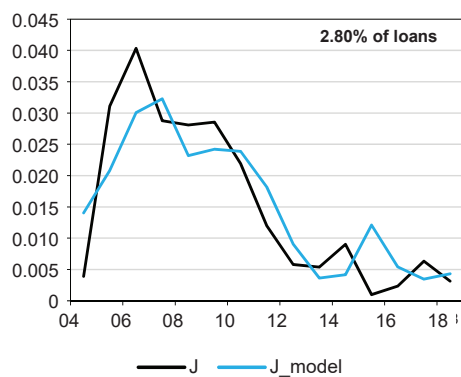
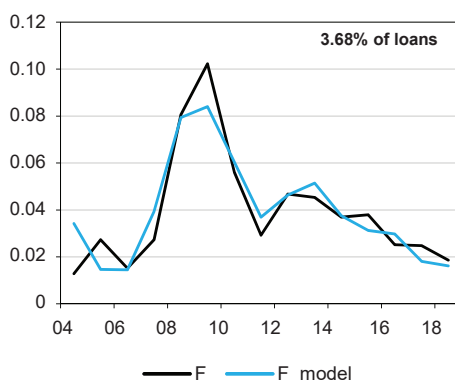
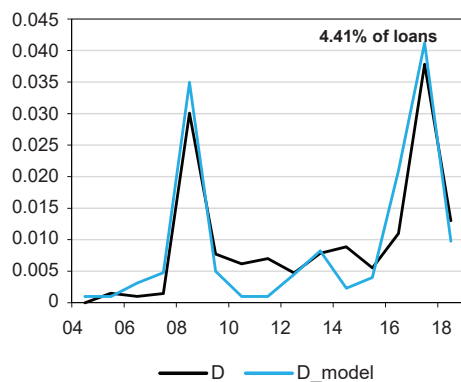
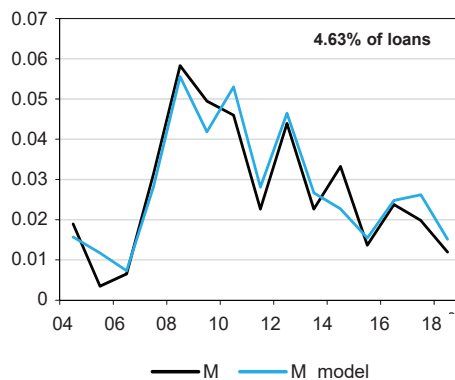
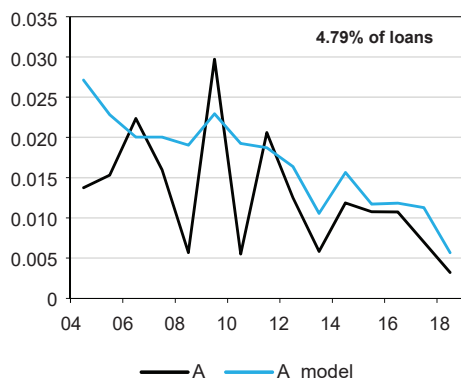
Note: A – Agriculture, forestry and fishing, B – Mining and quarrying, C – Manufacturing, D – Electricity, gas, steam and air conditioning supply, E – Water supply, sewerage and waste management, F – Construction, G – Wholesale and retail trade, repair of motor vehicles and motorcycles, H – Transportation 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 the learning process. Due to a very low signal-to-noise ratio, estimates for the sections A, H, L and N were performed on filtered series (Hodrick-Prescott filter with  $\lambda = 0.5$ ).

Source: CZSO, CNB, author's calculations

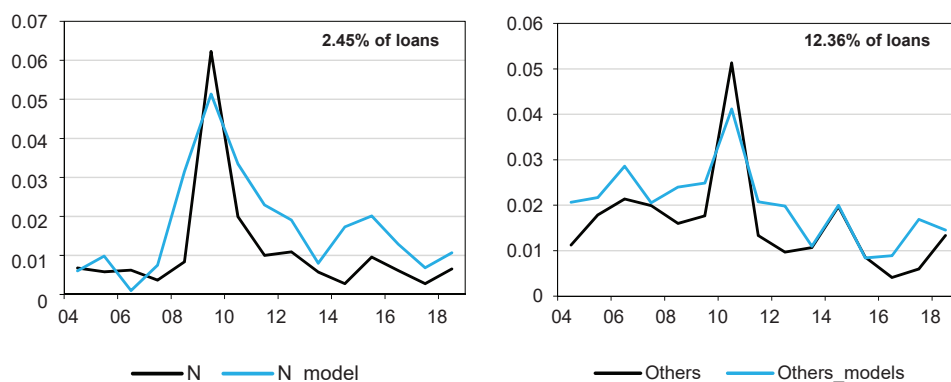
**Figure 3: Historical default rate estimates**



**Figure 3 (continuation)**



**Figure 3 (continuation)**



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 shares of individual sections in NFCs' total loan stock.

Source: CNB, author's calculations

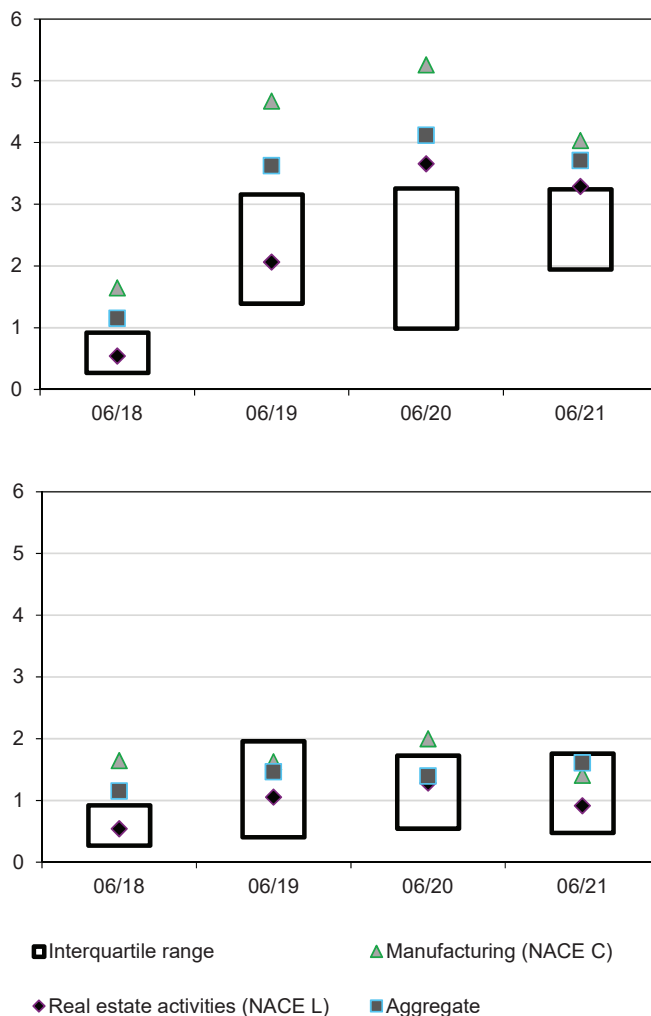
Figure 3 shows the default rate estimates for 2004–2018. The aggregate default rate is calculated as a weighted average of the sectional default rates. Due to low materiality, high noise in the time series and concentrated portfolios, sections with a share of total loans of less than 2% (5.6% of the 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 present 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.

### 3.2 Scenario A – Deep Recession

The first presented scenario is a standard pre-covid stress-test scenario. It 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 that originates from and is 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. It 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, the growth of corporate loans falls rapidly into negative figures. Private consumption shows greater resistance to the shock (losing 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.

For simplification, the shock to each component of GDP (final demand and imports) was proportionately divided into each economic division. Aggregate production in real terms falls to 95% of its maximum in the third year. Persistence in costs makes operating profits decrease sharply (the impacts on selected sections and divisions are shown in Figure A1). The presented scenario shows a significant rise in aggregate corporate defaults (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 longer recession than in 2009 and 2010, the forecasts show that non-financial corporations 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 manufacturing industry recorded the most significant contribution to the rise in the aggregate default rate, 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 4). On the contrary, the baseline scenario implies a moderate default rate for most sections. The suggested sum of defaulted loans is around 7% lower than in the Deep Recession scenario on a three-year horizon. The aggregate default rate is composed of the individual default rates of each division.

**Figure 4: Default rate forecasts in Deep Recession scenario and baseline scenario (in %)**



Note: 06/18 represents observed data.

Source: Observed data – CNB; forecasts – author’s calculations

### 3.3 Scenario B – COVID-19

As mentioned earlier, the global economy was being strongly affected by government measures to fight the spread of the COVID-19 pandemic at the time of writing. Given the rapid developments, one of the main problems during the COVID-19 crisis has

been a lack of relevant, live information. Forecasting the possible macroeconomic impacts of the COVID-19 shock in the early stage of the crisis motivated us to create a real-time COVID-19 scenario for the initial assessment of possible impacts on bank portfolios. 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 setting the initial controlling variables<sup>17</sup> 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). The averages of the core were used for the rest of the economy. 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 and demand nature 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 supply level 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).

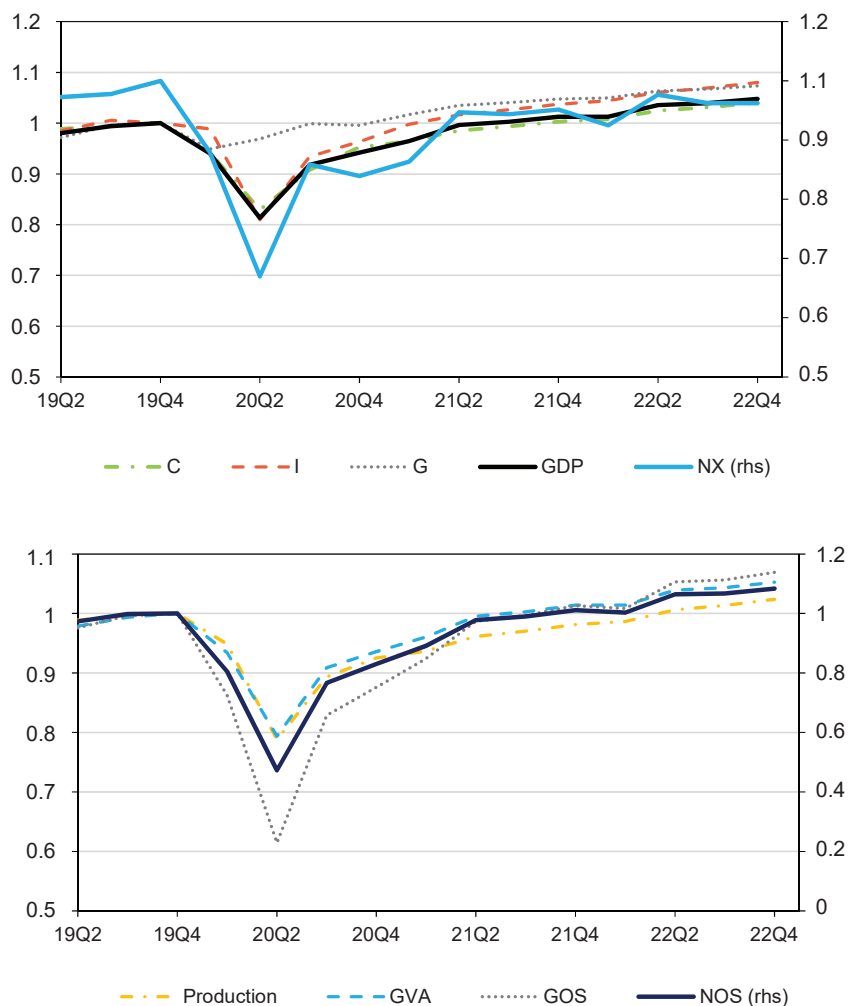
The presented scenario implied an immediate negative GDP growth (−2%) in 2020Q1, followed by a massive drop (17.1% YoY) in 2020Q2 and then a gradual recovery of the economy. Cumulatively, GDP falls by about 8.4% in 2020 in the COVID-19 scenario. For 2021 and 2022, the scenario implies convergence to 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. An increase

---

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

in property prices slows but does 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 5.

**Figure 5: 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 the liquidity need in the non-financial corporate sector. Considering the prescribed lockdown, the crisis arose “unnaturally” and strongly affected the financial health of all companies. There was a risk of an immediate slump in corporate cash flows, leading many firms with otherwise low credit risk to fall into insolvency and damaging the economy in the long run. Many governments introduced measures to overcome the imminent liquidity crisis to prevent job losses and a cascade of secondary insolvencies. A crucial policy question was how extensive the support should be. The simulation presented here gives a rough estimate of the minimum and sufficient liquidity support (Figure 6). The minimum liquidity support reflects the sum of the negative gross operating surpluses of the individual divisions of the economy. The sufficient support demonstrates the difference between gross operating surplus in the simulation and adjusted gross operating surplus in 2019Q4.<sup>18</sup>

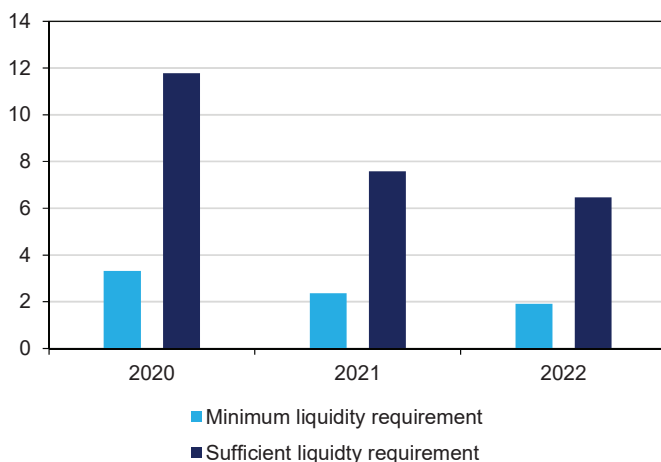
It was also presumed that the crisis would affect corporate credit quality and cause more loan defaults. The default rate forecasts were performed in two versions. Version A assumes no government support measures, while version B considers all the efforts announced up to 15 April 2020.<sup>19</sup> The aggregate default rates are shown in Figure 7.

---

18 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.

19 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 6: Minimum and sufficient liquidity support in COVID-19 scenario (% of nominal GDP in 2019)**

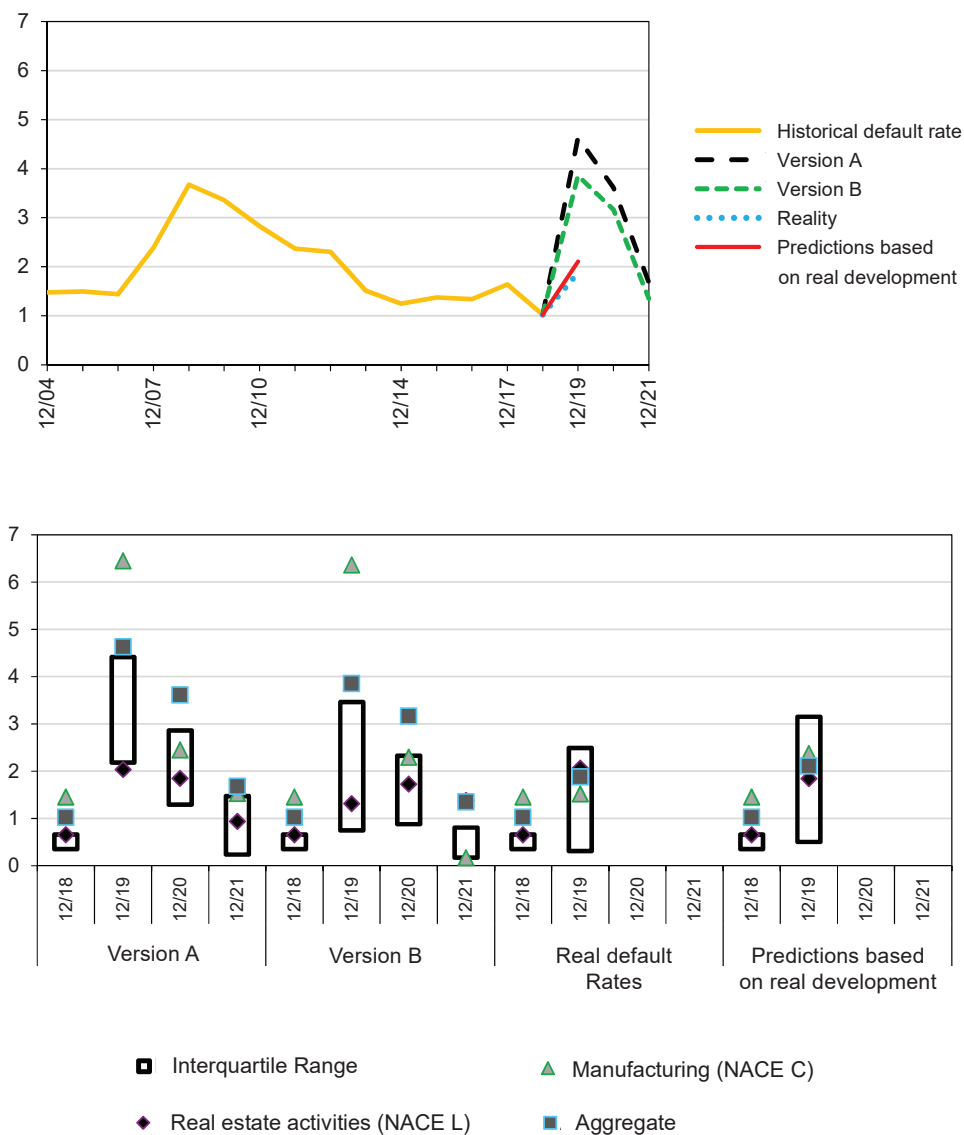


Source: Author's calculations

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 most significant in transportation and storage, construction and ICT. From a financial sector perspective, 8.4% of loan defaults could result in significant credit losses.

In November 2021, the forecasts for the first year of the prediction were evaluated. Real default rates show significantly lower realized credit risk than the results implied by the COVID-19 scenario. However, prediction of the default rate based on the observed section-specific and macroeconomic variables with coefficients presented in Section 4.1 (estimated for the period 2004-2018) is very accurate considering the unprecedented shock to the economy. The big difference between the April 2020 prediction and both observed data and the November 2021 prediction proves that the real-time constructed economic scenario was much more severe than the observed reality. A significant game changer was the size of government economic support and regulatory interventions, which went far beyond those considered and included in the scenario in April 2020. Nevertheless, these findings support the fact that the underlying scenario is crucial for default rate predictions and the hypothesis that estimated coefficients are robust.

**Figure 7: Default rates implied by COVID-19 scenario (%)**



Note: 06/18 represents observed data.

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

## Conclusion

This paper provides a framework for conducting simulations and stress testing in the non-financial corporate sector and shows possible outcomes, such as the sectional default rate predictions implied by the simulation results and the desired liquidity support during the current crisis. The present 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 emission 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.

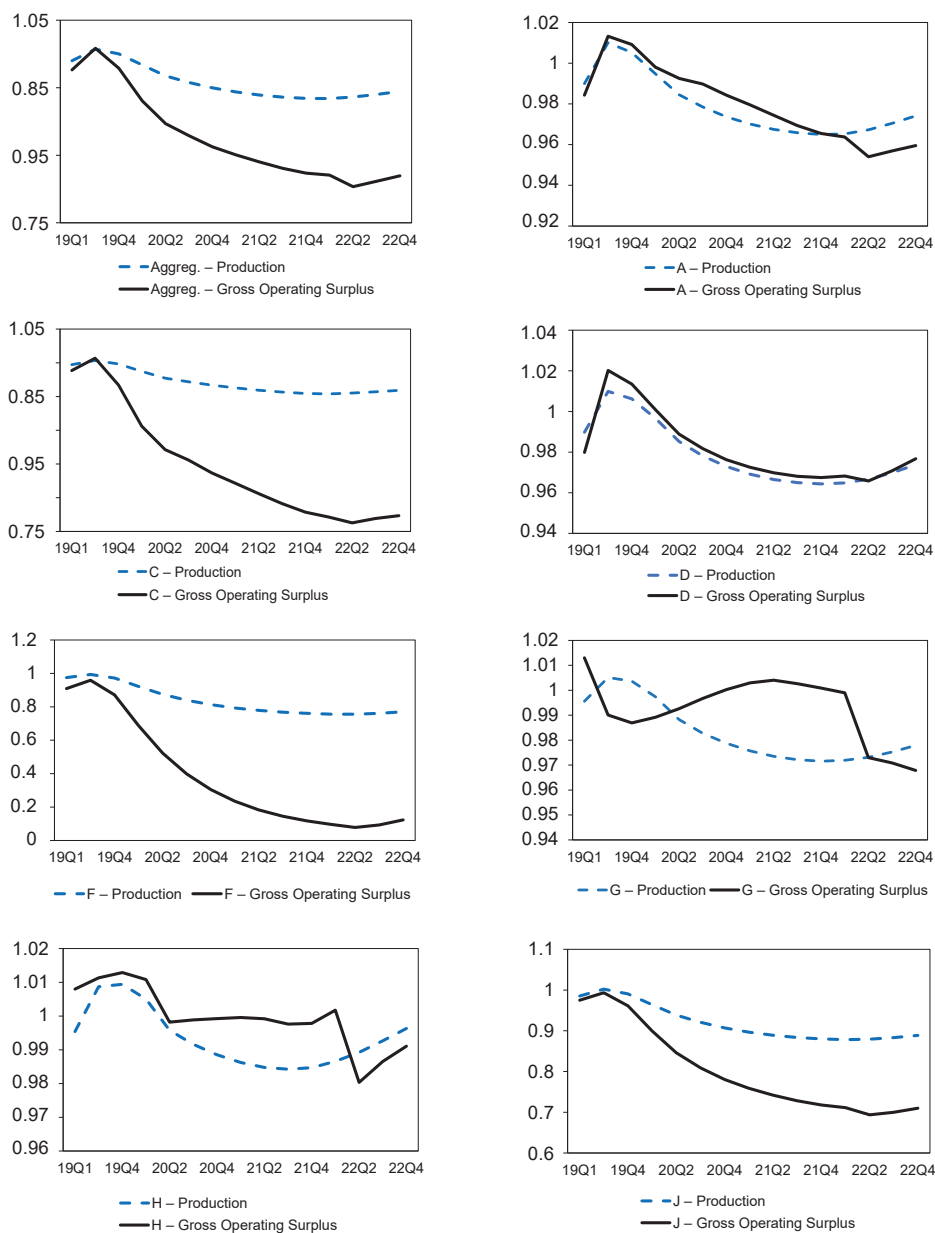
The underlying scenario is strictly exogenous in the present approach and can be taken from fiscal, monetary or international institutions, as they perform detailed economic predictions on a regular basis. Results of the proposed approach strongly depend on the underlying scenario. The results are scenario-conditional coherent projections of economic performance and credit risk of individual NACE sections of the corporate sector. Scenario dependency is shown in Section 4.3, where it was created flexibly in real time at the beginning of the COVID-19 crisis. Default rates implied by this scenario were far higher than the real observed default rates. The high difference was caused by the scenario being much more severe than reality, considering the accurate default rate forecast based on real values.

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 use of other approaches to estimate credit risk coefficients may be beneficial, such as BMA methods etc. (see Panoš and Polák, 2020).

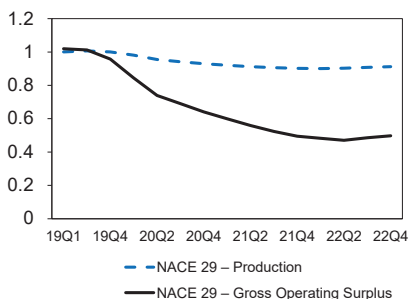
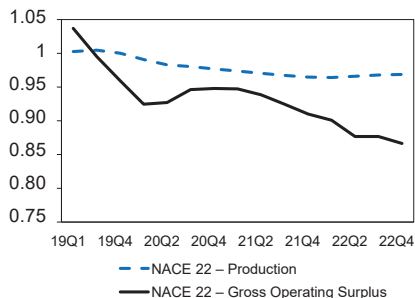
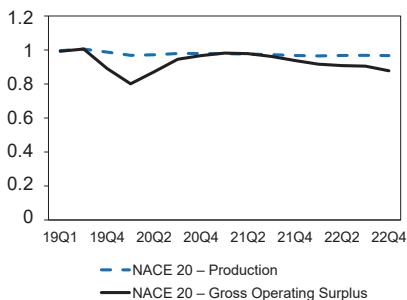
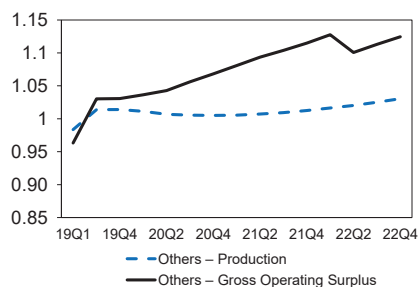
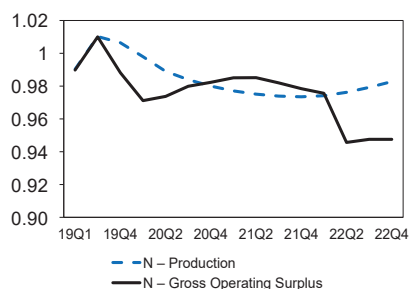
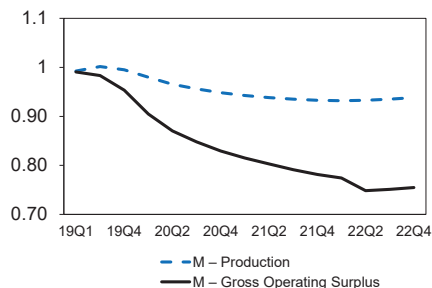
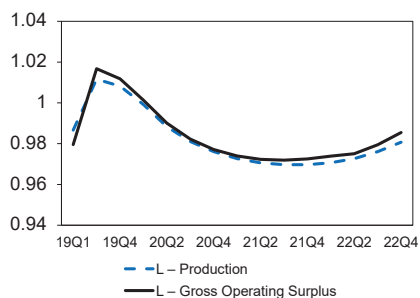
The results also contain reliable historical estimates of corporate default rates for the main sections of the Czech economy in the period 2004–2018. Finally, other outcomes presented in the paper, including the range of additional liquidity support needed and an evaluation of the measures implemented, contribute to the policy debate.

## Appendix

**Figure A1: Production and gross operating surplus of selected sections and divisions in Deep Recession scenario (real terms index: 2019Q2 = 1)**



**Figure A1 (Continuation)**



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.

Source: Autor's own calculation

**Figure A2: Algorithm adjustment for capturing supply-side shock (from R)**

```

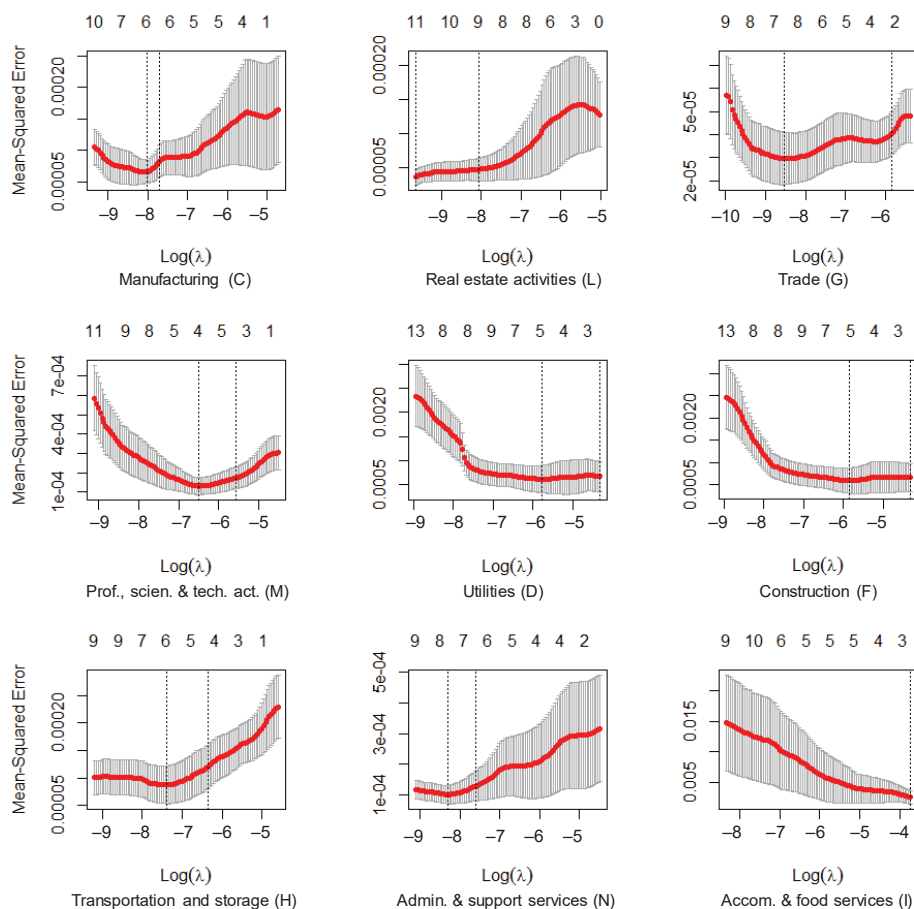
if (d_s(t) > d_sf(t))
{d_s(t) <- d_sf(t)
  Intermediate_Cons(t) <- Intermediate_Cons(t-1) + (d_s(t))*A
  Final_demand(t) <- Final_demand(t-1) + (d_s(t))*(1-rowSums(A))
}

```

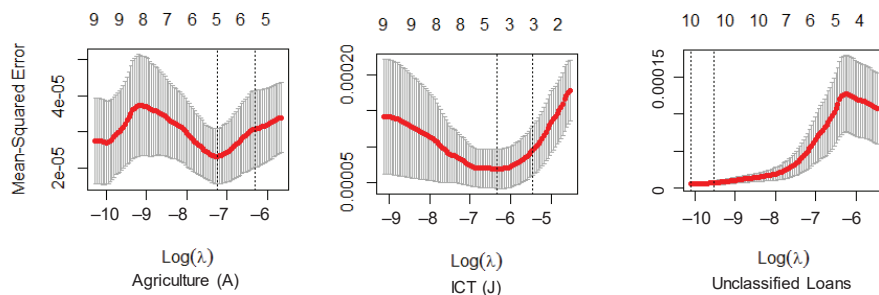
Note:  $d_s(t)$  stands for change in supply implied by demand shock,  $d_{sf}(t)$  represents the primary assumption about the supply shock, and  $A$  is the technology matrix (defined on page 3).

Source: Autor's own elaboration

**Figure A3: Cross-validation characteristics for major sections**



**Figure A3 (Continuation)**



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

Source: Autor's own elaboration

**Figure A4: Lambda traceplots for all sections**

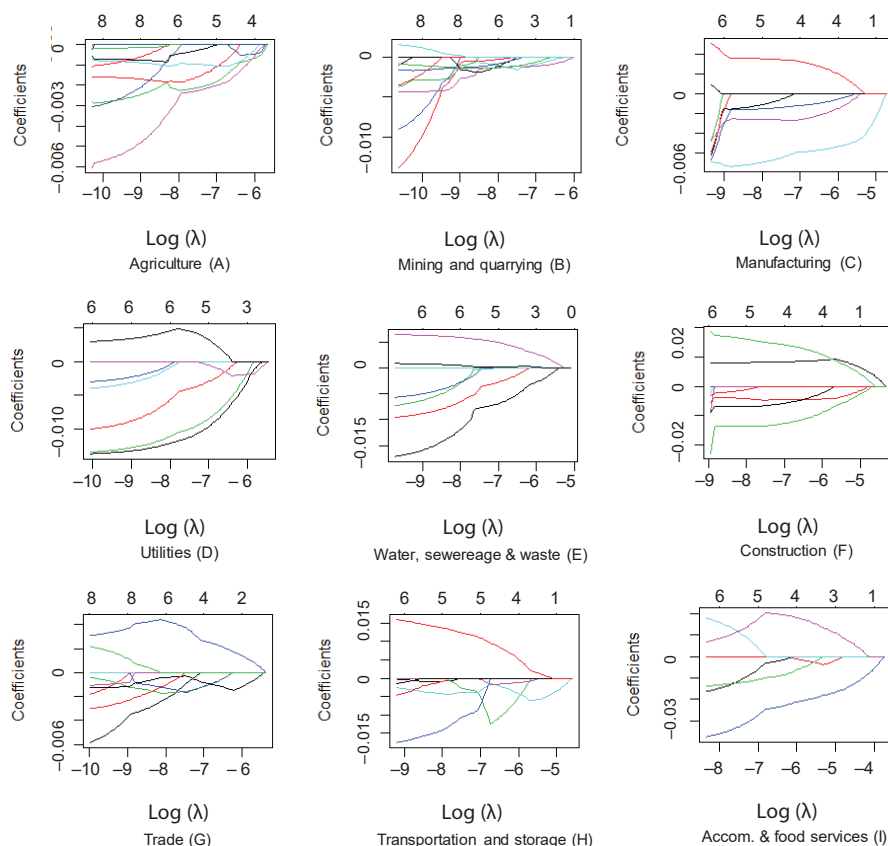
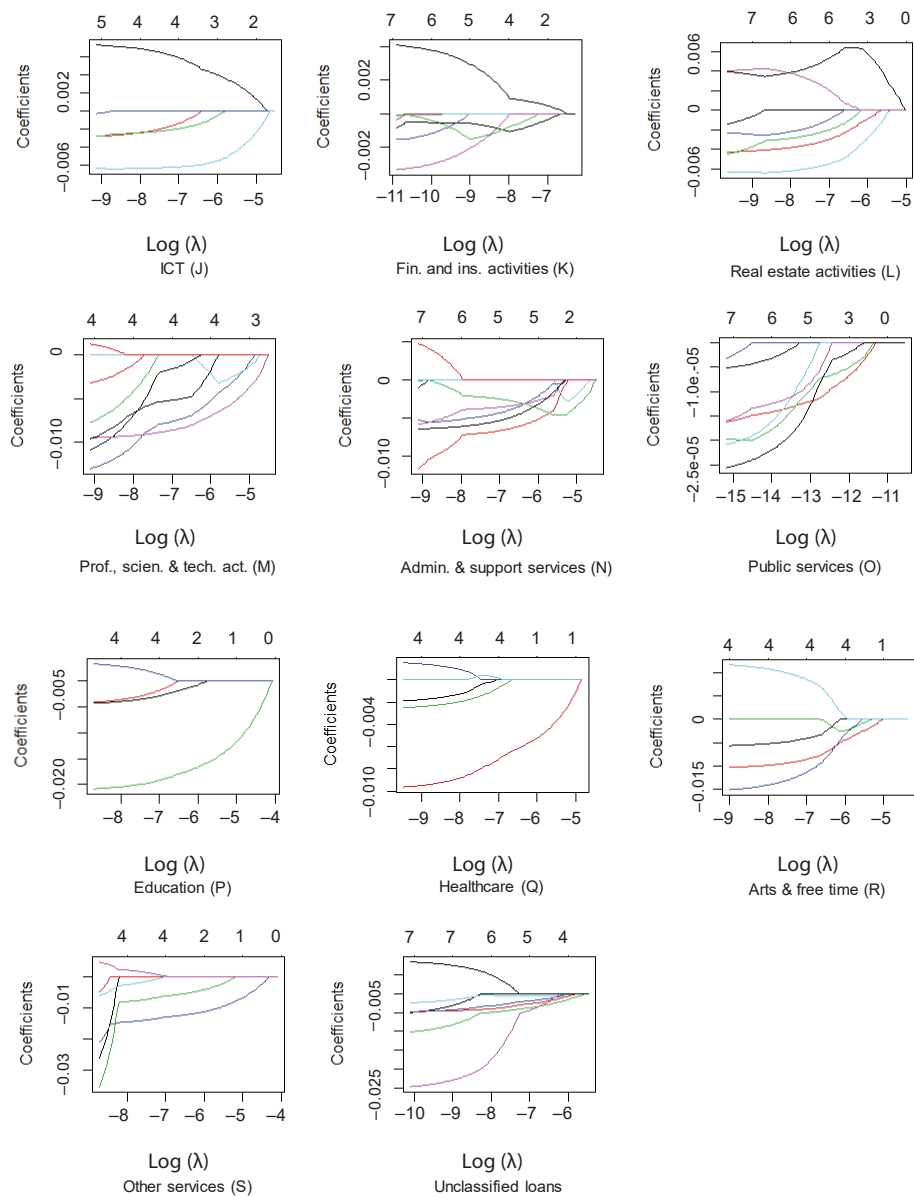


Figure A4 (Continuation)



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

Source: Autor’s own elaboration

## References

- Agarwal, V., Taffler, R. (2008). Comparing the Performance of Market-based and Accounting-based Bankruptcy Prediction Models. *Journal of Banking & Finance*, 32(8), 1541–1551, <https://doi.org/10.1016/j.jbankfin.2007.07.014>
- Alatríste Contreras, M., Fagiolo G. (2014). *Propagation of Economic Shocks in Input-Output Networks*. LEM Working Paper Series No. 2014/09. Available at: <https://arxiv.org/pdf/1401.4704.pdf>
- Baum, C. F., Caglayan, M., Barkoulas, J. T. (2001). Exchange Rate Uncertainty and Firm Profitability. *Journal of Macroeconomics*, 23(4), 565–576, [https://doi.org/10.1016/S0164-0704\(01\)00178-1](https://doi.org/10.1016/S0164-0704(01)00178-1)
- Blackwood, G. J. (2018). Same Shock, Separate Channels: House Prices and Firm Performance in the Great Recession. Amherst College Working Paper Series. Available at: [https://www.amherst.edu/system/files/JMP\\_Housing\\_v5\\_1.pdf](https://www.amherst.edu/system/files/JMP_Housing_v5_1.pdf)
- BCBS (2000). *Principles for the Management of Credit Risk*. Basel Committee on Banking Supervision. Available at: <https://www.bis.org/publ/bcbs75.pdf>
- Castrén, O., Dées, S., Zaher, F. (2009). Stress-testing Euro Area Corporate Default Probabilities Using a Global Macroeconomic Model. *Journal of Financial Stability*, 6(2), 64–78, <https://doi.org/10.1016/j.jfs.2009.10.001>
- Cipollini, A., Missaglia, G. (2008). Forecasting Industry Sector Default Rates Through Dynamic Factor Models. *Journal of Risk Model Validation*, 2(3), 59–79, <https://doi.org/10.21314/JRMV.2008.027>
- Dixon, P., Jorgenson, D. W. (2013). *Handbook of Computable General Equilibrium Modeling*. Amsterdam: Elsevier. ISBN: 9780444536341.
- Drehmann, M. (2005). A Market Based Macro Stress Test for the Corporate Credit Exposures of UK Banks. London: Bank of England, <https://doi.org/10.2139/ssrn.676850>
- Eurostat (2006). *NACE Rev. 2 – Statistical Classification of Economic Activities*. Luxembourg: Eurostat. ISBN: 978-92-79-04741-1.
- Feyen, E., Fiess, N., Zuccardi Huertas, I., et al. (2017). *Which Emerging Markets and Developing Economies Face Corporate Balance Sheet Vulnerabilities?* Policy Research Working Paper Series No. 8198, World Bank. Available at: <https://openknowledge.worldbank.org/bitstream/handle/10986/28376/WPS8198.pdf?sequence=1&isAllowed=y>
- Figlewski, S., Frydman, H., Liang, W. (2012). Modeling the Effect of Macroeconomic Factors on Corporate Default and Credit Rating Transitions. *International Review of Economics and Finance*, 21(1), 87–105, <https://doi.org/10.1016/j.iref.2011.05.004>
- Flom, P. (2018). Stopping Stepwise: Why Stepwise Selection Is Bad and What You Should Use Instead. In: *Towards Data Science*. Available at: <https://towardsdatascience.com/stopping-stepwise-why-stepwise-selection-is-bad-and-what-you-should-use-instead-90818b3f52df>

- Frait, J., Komárková, Z. (2012). *Macroprudential Policy and Its Instruments in a Small EU Economy*. CNB Research and Policy Notes No. 3/2012. Available at: [https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research\\_publications/irpn/download/rpn\\_3\\_2012.pdf](https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research_publications/irpn/download/rpn_3_2012.pdf)
- Friedman, J. H., Hastie, T., Tibshirani, R. (2016). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33(1), 1–22, <https://doi.org/10.18637/jss.v033.i01>
- Geršl, A., Jakubík, P., Konečný, T., et al. (2013). Dynamic Stress Testing: The Framework for Assessing the Resilience of the Banking Sector Used by the Czech National Bank. *Czech Journal of Economics and Finance (Finance a úvěr)*, 63(6), 505–536.
- Harrell, F., E. (2001). *Regression Modeling Strategies with Applications to Linear Models, Logistic Regression and Survival Analysis*. New York: Springer-Verlag. ISBN: 978-3-319-19425-7.
- Hohn, K. (2004). Application to Economics: Leontief Model. Manhattan: Kansas State University. Available at: <https://www.math.ksu.edu/~gerald/leontief.pdf>
- Jakubík, P. (2007). Macroeconomic Environment and Credit Risk. *Czech Journal of Economics and Finance (Finance a úvěr)*, 57(1–2), 60–78. Available at: [https://journal.fsv.cuni.cz/storage/5\\_fau\\_1\\_2\\_07\\_00000000060.pdf](https://journal.fsv.cuni.cz/storage/5_fau_1_2_07_00000000060.pdf)
- Jakubík, P., Schmieder, C. (2008). *Stress Testing Credit Risk: Is the Czech Republic Different from Germany?* CNB Working Papers, 9/2008. Available at: [https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research\\_publications/cnb\\_wp/cnbwp\\_2008\\_09.pdf](https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research_publications/cnb_wp/cnbwp_2008_09.pdf)
- Kadlčáková, N., Keplinger, J. (2004). *Credit Risk and Bank Lending in the Czech Republic*. CNB Working Papers No. 6/2004. Available at: [https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research\\_publications/cnb\\_wp/cnbwp62004.pdf](https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research_publications/cnb_wp/cnbwp62004.pdf)
- Karasulu, M., Jones, M. T. (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 No. 06/114, <https://doi.org/10.5089/9781451863741.001>
- Klein, N., Murgašová, Z. (2016). *Corporate Sector Vulnerabilities in Ireland*. IMF Working Papers No. 16/211, <https://doi.org/10.5089/978147552171.001>
- Lee, K.-M., Yang, J.-S., Kim, G., et al. (2011). Impact of the Topology of Global Macroeconomic Network on the Spreading of Economic Crises. *PLoS ONE*, 6(3), <https://doi.org/10.1371/journal.pone.0018443>
- Leontief, W. (1936). Quantitative Input and Output Relations in the Economic System of the United States. *Review of Economics and Statistics*, 18(3), 105–125, <https://doi.org/10.2307/1927837>
- Lequiller, F., Blades, D. (2014). *Understanding National Accounts: Second Edition*. Paris: OECD Publishing. ISBN: 978-92-64-21463-7.
- Louzis, D. P., Vouldis, A. T., Metaxas, V. L. (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(4), 1012–1027, <https://doi.org/10.1016/j.jbankfin.2011.10.012>

- Panoš, J., Polák, P. (2020). *How to Improve the Model Selection Procedure in a Stress-testing Framework*. CNB Working Papers No. 9/2019. Available at: [https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research\\_publications/cnb\\_wp/cnbwp\\_2019\\_09.pdf](https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research_publications/cnb_wp/cnbwp_2019_09.pdf)
- Radivojević, N., Jovović, J. (2017). Examining of Determinants of Non-Performing Loans. *Prague Economic Papers*, 26(3), 300–316, <https://doi.org/10.18267/j.pep.615>
- Simons, D., Rolwes, F. (2009), Macroeconomic Default Modelling and Stress Testing. *International Journal of Central Banking*, 5(3), 177–204. Available at: <https://www.ijcb.org/journal/ijcb09q3a6.pdf>
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Virolainen, K. (2004). Macro Stress Testing with a Macroeconomic Credit Risk Model for Finland. Bank of Finland Discussion Papers No. 18/2004, <https://doi.org/10.2139/ssrn.622682>