Indeks 371262
e-ISSN 2657-9545
ISSN 0033-2372

Przegląd Statystyczny
Statistical Review

Vol. 69 | No. 2 | 2022

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

INFORMATION FOR AUTHORS

Przegląd Statystyczny. Statistical Review publishes original research papers on theoretical and empirical topics in statistics, econometrics, mathematical economics, operational research, decision science and data analysis. The manuscripts considered for publication should significantly contribute to the theoretical aspects of the aforementioned fields or shed new light on the practical applications of these aspects. Manuscripts reporting important results of research projects are particularly welcome. Review papers, shorter papers reporting on major conferences in the field, and reviews of seminal monographs are eligible for submission, but only on the Editor's request.

Since 1 May 2019, the journal has been publishing articles in English.

Any spelling style is acceptable as long as it is consistent within the manuscript.

All work should be submitted to the journal through the ICI Publishers Panel (https://editors.publisherspanel.com/pl.ici.ppanel-app-war/ppanel/index).

For details of the submission process and editorial requirements please visit https://ps.stat.gov.pl/ForAuthors.

Przegląd Statystyczny
Statistical Review

Vol. 69 No. 2 2022

EDITORIAL OFFICE'S ADDRESS

University of Gdańsk, ul. Armii Krajowej 101, 81-824 Sopot, Poland

Language editing: Scientific Journals Division, Statistics Poland
Technical editing and typesetting: Statistical Publishing Establishment – team supervised by Wojciech Szuchta

Zakład Wydawnictw Statystycznych

Printed and bound: Statistical Publishing Establishment
al. Niepodległości 208, 00-925 Warsaw, Poland, zws.stat.gov.pl

Website: ps.stat.gov.pl

ISSN 0033-2372
e-ISSN 2657-9545
Index 371262

Information on the sales of the journal: Statistical Publishing Establishment
phone no. +48 22 608 32 10, +48 22 608 38 10

Order no. 255/2022 – 210 printed copies

CONTENTS

Przegląd Statystyczny. Statistical Review, 2022, vol. 69, 2, 1–19 DOI: 10.5604/01.3001.0016.0363
Received: 12.07.2022, revised: 12.09.2022, accepted: 25.09.2022

Do mixed-data sampling models help forecast liquidity and volatility?

Barbara Będowska-Sójka,[a] Agata Kliber[b]

Abstract. This paper aims to contribute to the existing studies on the Granger-causal relationship between volatility and liquidity in the stock market. We examine whether liquidity improves volatility forecasts and whether volatility allows the improvement of liquidity forecasts. The forecasts based on the mixed-data sampling models, MIDAS, are compared to those obtained from models based on daily data. Our results show that volatility and liquidity forecasts from MIDAS models outperform naive forecasts. On the other hand, the application of mixed-data sampling models does not significantly improve the performance of the forecasts of either liquidity or volatility based on a univariate autoregressive model or a vector-autoregressive one. We found that in terms of the forecasting ability, the VAR models and the AR models seem to perform equally well, as the differences in forecasting errors generated by these two types of models are not statistically significant.

Keywords: liquidity, volatility, effective spread estimator, MIDAS
JEL: G12, G15

1. Introduction

Volatility and liquidity of the financial instruments are the core concepts in empirical finance. The first is usually defined as the statistical measure of the dispersion of returns for a given security, while the second is described as the ability to buy or sell an asset immediately at a low cost without affecting the asset's price significantly (Pástor & Stambaugh, 2003). Volatility and liquidity share some common features: both are unobservable, difficult to estimate and time-varying. There is no simple answer to the question what the best proxy for either volatility (Andersen et al., 2007) or liquidity (Díaz & Escribano, 2020) is. Here two approaches are commonly applied: volatility and liquidity measures are based either on data of the same frequency (e.g. daily measures based on daily data) or on data of higher frequency (e.g. daily measures based on intradaily data) (Ahn et al., 2018; Andersen & Bollerslev, 1998). Generally, measures based on higher-frequency data

[a] Poznań University of Economics and Business, Institute of Informatics and Quantitative Economics, Department of Econometrics, al. Niepodległości 10, 61-875 Poznań,
e-mail: barbara.bedowska-sojka@ue.poznan.pl, ORCID: https://orcid.org/0000-0001-5193-8304.
[b] Poznań University of Economics and Business, Institute of Informatics and Quantitative Economics, Department of Applied Mathematics, al. Niepodległości 10, 61-875 Poznań,
e-mail: agata.kliber@ue.poznan.pl, ORCID: https://orcid.org/0000-0003-1996-5550.

should be more informative, as the set of information is more comprehensive (Giot, 2005). However, such data are usually expensive and therefore not available for all investors. The time-varying feature was exhaustively examined both in the case of volatility (Faff et al., 2000) and liquidity (Liang & Wei, 2012). As such, these variables are also difficult to predict.

The aim of the paper is to examine two issues. Firstly, we investigate whether information on the past liquidity can improve volatility forecasts, and vice versa – whether data on previous volatility can improve liquidity forecasts. Secondly, we consider the application of mixed-frequency data by comparing the accuracy of forecasts from mixed-data sampling models, MIDAS (Ghysels et al., 2004), to those which use variables in one frequency only. For the latter, we consider vector autoregressive (VAR) models with the other variable as the regressor, and simple autoregressive (AR) models without any additional variables. We examine which method, the one employing mixed-frequency data or the one applying one-frequency data only, generates better results in terms of out-of-sample volatility and liquidity forecasts.

We employed a dataset from the European emerging market, the Warsaw Stock Exchange (further: the WSE), which was a sample of 118 stocks listed on this market and observed over a period of eight years. Such a forecasting exercise requires liquidity and volatility measures that could be obtained for a low (daily) and high (intradaily) frequencies. Thus, volatility in our approach was approximated by a realised variance (Andersen et al. 2006), while liquidity was calculated as the quoted effective spread of Chung and Zhang (2014). The former measure is identified in the literature as a good proxy for volatility (Andersen & Bollerslev, 1998), and the latter for liquidity (Fong et al., 2017; Ma et al., 2018).

The main result of the study was finding no advantage in using MIDAS models. Models based on daily data only, such as univariate AR or bi-variate VAR ones, performed better than the more complicated AR_MIDAS ones, where mixed frequencies were applied. There was no distinction between the AR and the VAR models – both were performing equally well within the forecasting framework. Among the specifications considered, the MIDAS model outperformed only the naive approach.

The remaining part of the paper is organised in the following way: Section 2 is devoted to the literature review on the dependency between volatility and liquidity, Section 3 describes the sample and variables used in the study, Section 4 shows the research methodology, Section 5 presents empirical results, and Section 6 summarises and concludes the study.

2. Literature review

The literature shows that volatility and liquidity are interrelated. Chordia et al. (2001) found that aggregated liquidity is influenced by recent market volatility, among other factors. Rösch and Kaserer (2014) showed that liquidity increases in the time of market downturns, while Yeyati et al. (2008) described the 'spiralling fall' effect, which manifests itself in lower liquidity when stock market returns decrease rapidly and volatility is higher. The faster the market falls, the less liquidity there is. Brunnermeier and Pedersen (2009) showed that higher volatility tends to increase illiquidity, because financial intermediaries reduce their activity in volatile times. Ma et al. (2018) found the dependence between stock market volatility and trading activity, namely as the market becomes more volatile, the trading volume decreases.

Another current in the literature focuses on the causal relationships between liquidity and volatility. There is evidence for a one-direction or bi-directional causality in different stock markets (Będowska-Sójka & Kliber, 2019; Hautsch & Jeleskovic, 2008; Hiemstra & Jones, 1994; Gold et al., 2017). According to the causality definition, if one time series is a Granger cause for another, it improves the latter's forecasts (Ong, 2015). Therefore it seems that combining volatility and liquidity in the forecasting framework and using one of them when predicting the other might be effective.

This study is the extension of the previous research by Będowska-Sójka and Kliber (2019). That former research also pertained to the WSE and showed that there was a causal relationship between volatility and liquidity and vice versa. Moreover, it was demonstrated that liquidity reacted differently to the increase and the decrease in volatility, and likewise volatility – it was affected to a different extent by the rise and the decline in liquidity. A natural extension of that study would be to find out whether the causal relationships are strong enough to be useful in forecasting. Moreover, intraday data seems to be more relevant, as it brings more information about the market than the daily data. Here a question arises whether additional information is useful in predicting the aforementioned measures.

As already mentioned, to address the above issues, we first estimated and then generated volatility and liquidity forecasts from the following models: the MIDAS, the VAR and the AR. In the literature, the MIDAS model is successfully used to describe the dynamics of macroeconomic variables. For instance, Smith (2016) and Maas (2019) used the MIDAS model to successfully nowcast the unemployment by means of Google-search data as a high-frequency regressor. There is evidence that the MIDAS regression outperforms other models in predicting GDP (Ferrara & Marsilli, 2013; Kim & Swanson, 2018; Tsui et al., 2018). Also Andreou et al. (2010) found that using regressions with differently-sampled data improved the forecasting ability of the empirical economic growth. Other authors showed that incorporating mixed-frequency data to inflation modelling had promising results. Breitung and Roling (2015) demonstrated that the commodity price index is a useful predictor of inflation rates 20–30 days ahead, and Monteforte and Moretti (2013) found that the inclusion of daily variables from the financial market in the model of monthly inflation helps to reduce forecast errors.

Many researchers also proved that the MIDAS model could be successfully applied to both modelling and forecasting of the financial-market data. Although in the MIDAS-GARCH approach (Engle et al., 2013), the high-frequency volatility is modelled with low-frequency data, as e.g. economic indicators (Asgharian et al., 2013; Engle et al., 2013) or other regressors of lower frequency (Ma et al., 2019), there have also been attempts to model the daily volatility with intradaily data. Such a mixed-data sampling approach was applied to volatility prediction by Ghysels et al. (2006), who used high- and low-frequency data to forecast volatility. Their model allowed the improvement of forecasts by 30% compared to the benchmark model. Further, Santos and Ziegelmann (2014) juxtaposed several multi-period volatility forecasting models from the MIDAS and the HAR families in order to forecast the future volatility of the BOVESPA index. They concluded that regressors involving volatility measures robust to jumps are better in forecasting the future volatility – which corroborates the findings described in Ghysels et al. (2006) – and that the relative forecasting performances of the three approaches are comparable.

To our best knowledge, there have not been so far any such attempts when liquidity and volatility were forecasted. There is still no evidence whether the incorporation of high-frequency measures of volatility (or liquidity) is helpful when forecasting liquidity (or volatility) in daily frequency. The presence of causality between volatility and liquidity justifies such experiments.

3. The description of the dataset used in the study

3.1. Data source and sample description

The sample duration extended from January 2009 until December 2016. The stocks included in the sample were constantly listed on the WSE throughout this period. The final sample consisted of 118 stocks well-established in the market and with a relatively long history of listing (the full list is available from the corresponding author upon request). We used high-frequency data from the database constructed on the basis of data offered directly by the WSE.

As the original source were tick-by-tick data, they had to be pre-processed. The procedures described in Barndorff-Nielsen et al. (2009) were applied, which made it possible to control for outliers, multiple or missing records, and other incidents that might occur in high-frequency datasets. Then the filtered tick-by-tick data were aggregated into equally-sampled 10-minute, 30-minute, and 60-minute data. Thus we received eight years of data for 118 stocks with four different frequencies: three intraday and one daily.

3.2. Volatility and liquidity proxies

As volatility and liquidity are unobservable, we used non-parametric measures to calculate the daily and intradaily estimates. The choice of the proxies was based on the fact they were relatively easy to calculate and it was possible to obtain the estimates in different frequencies: daily and intradaily. Volatility was proxied by realised variance (RV), and calculated as (Andersen et al., 2007):

$$RV_t = \sum_{i=1}^{I} r_{i,t}^2, \tag{1}$$

where RV_t is a daily realised variance in day t, r_{it} is a log return in interval i (e.g. 10-minute), and I is the number of intra-daily periods within a day. The realised variance is one of the estimators of volatility that are most frequently used (Andersen & Bollerslev, 1998; Fuertes & Olmo, 2012; Laurent & Violante, 2012).

Out of various liquidity proxies, we chose the closing quoted spread, CQS, of Chung and Zhang (2014). The following formula was applied:

$$CQS_t = \frac{A_t - B_t}{0.5(A_t + B_t)}, \tag{2}$$

where B_t and A_t were the bid and the ask prices, respectively, at the end of a given day *t*. Various studies showed that the CQS is the best proxy for unobserved liquidity (Chung & Zhang, 2014; Díaz & Escribano, 2020; Fong et al., 2017).

We also calculated these two measures in the high-frequency setting: the realised volatility were the squares of intradaily returns in a given sampling frequency, while the quoted effective spread was calculated on the basis of the last bid and ask price within a given time interval (e.g. 1 hour).

4. Methodology

4.1. The MIDAS model

We used the following notation: y_t was a dependent variable representing a low-frequency process and sampled at daily frequency while x_τ was an explanatory variable sampled in high frequency. For x_τ, we considered three distinct cases: a 10-minute, a 30-minute, and a 60-minute frequency. Please note that $By_t = y_{t-1}$ and $Lx_\tau = x_{\tau-1}$ were the lags of the low-frequency and the high-frequency processes, respectively. It was assumed that for each low-frequency period $t = t_0$, we observed high frequency process $x_\tau^{(i)}$ at $m_i \in N$ intervals: $\tau = (t_0 - 1)m_i + j$, $j = 1, \dots, m_i$.

Since the session schedule within our sample period changed three times, we choose to consider records from 9.00 a.m. to 4 p.m. Due to some irregularities in the data, and in order to conveniently define equally sampled series, we had to skip the first observation, when data was sampled at the frequency higher than 1 hour. For 10-minute data, the first observation was made at 9.10 a.m., while the last one was recorded at 4 p.m. Thus we have $m_1 = 42$ observations of the high-frequency process, and $\tau = 0, \dots, 42$. When x was sampled at a 30-minute frequency, the first observation came at 9.30 a.m., while the last one was made at 4 p.m., thus: $m_2 = 14$ and $\tau = 0, \dots, 14$. Finally, when x was sampled every 60 minutes, the first observation came at 9:00 a.m., while the last one was made at 4 p.m., so $m_3 = 8$ and $= 0, \dots, 8$. In each of the above cases, there was only one observation per day for the low-frequency process.

The MIDAS model can be written in a compact form as (Ghysels et al., 2016):

$$\alpha(B)y_t = \beta(L)^T \mathbf{x}_{t,0} + \epsilon_t, \tag{3}$$

where:
$\alpha(z) = 1 - \sum_{j=1}^{p} \alpha_j z^j$ (low-frequency lag operator),
$\mathbf{x}_{t,0} := \left(x_{tm_0}^{(0)}, \dots, x_{tm_i}^{(i)}, \dots, x_{tm_l}^{(l)}\right)^T$,
$\beta(z) = 1 - \sum_{j=1}^{p} \beta_j z^j$ (high-frequency lag operator), $\beta_j = \left(\beta_j^{(0)}, \dots \beta_j^{(i)}, \beta_j^{(l)}\right)$,
T denotes transposition, and i the frequency period.

In our study, we considered AR(1)-MIDAS models, and in each model we included explanatory variables of only one frequency (either 10-minute, 30-minute, or 60-minute). Thus, our model can be specified in an alternative form as:

$$y_t = ay_{t-1} + \sum_{j=0}^{l} \beta_j x_{tm-j} + \epsilon_t\,. \tag{4}$$

To estimate the model, one needs to align the high-frequency data to the low-frequency data. The alignment is performed through the following transformation (Ghysels et al., 2016):

$$\sum_{j=0}^{l} \beta_j x_{tm-j} = \sum_{r=0}^{q} \lambda_r \tilde{x}_{t-r} \,, \tag{5}$$

where $q \in N$ denotes a low-frequency number of lags, and $\tilde{x}_{t-r}$ the parameter-driven low-frequency aggregates (Ghysels et al., 2016):

$$\tilde{x}_{t-r} = x_{t-r}(\boldsymbol{\delta}_r) = \sum_{s=1}^{m} \omega_r(\boldsymbol{\delta}_r; s) x_{(t-r-1)m+s} \,, \tag{6}$$

The function $\omega_r(\boldsymbol{\delta}_r; s)$ is called a weighting function, and its parameter vector $\boldsymbol{\delta}_r$ can generally vary with each variable and low-frequency lag order r. The aggregation weights λ_r are usually non-negative and satisfy the normalisation constraints: $\sum_{s=0}^{m-1} \omega_r(\boldsymbol{\delta}_r; s) = 1$. To have the weights add to one, it is convenient to define a weighting function in the following form (Ghysels et al., 2016):

$$\forall r: \omega_r(\boldsymbol{\delta}_r; s) = \frac{\psi_r \omega_r(\boldsymbol{\delta}_r; s)}{\sum_{j=1}^{m} \psi_r(\boldsymbol{\delta}_r; j)}, \qquad s = 1, \dots, m, \tag{7}$$

where $\psi_r(\cdot)$ denotes some underlying function. If the latter is non-negatively valued and the denominator is positive, the weights (7) are also non-negative (Ghysels et al., 2016).

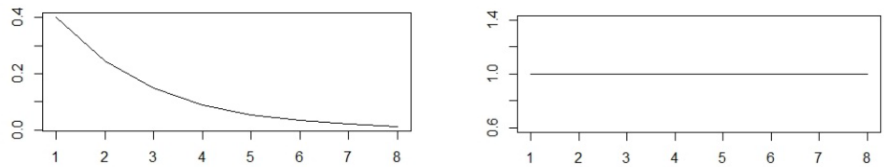
There are various possible specifications of the underlying functions described in the literature: an exponential Almon lag polynomial, beta function, Gomperts, log--Cauchy, etc. (see Ghysels et al., 2016 for details). Using the constraint function has two advantages. Firstly, it allows the reduction of the number of parameters in the model. Secondly, if the parameters of an underlying data-generating process follow a certain functional constraint, and this constraint is well-approximated by a chosen constraint function, the accuracy of the out-of-sample predictions can improve significantly – as shown by Ghysels et al. (2016).

In our study, we use the exponential Almon lag polynomial:

$$\psi(\boldsymbol{\delta}; s) = \exp\left(\sum_{j=1}^{p} \delta_j s^j\right), p \in N, \tag{8}$$

in its normalized and non-normalized form. The Almond polynomial is flexible and can take various shapes with only a few parameters (Ghysels et al., 2007). As a starting point for the estimation, we parametrised the constraint function in such a way that the newest observations had higher weights than the older ones (Almon function with two parameters: 1 and –0.5). As an alternative, equal weights were tested. We present the lag functions in Figure 1.

Figure 1. Alternative shapes of weight functions used in the MIDAS model specification



Source: authors' calculations.

There are several phases of a model selection. In each of them, we took into consideration all the information from the day (i.e. all eight intra-daily observations of the regressor in the hourly model, and 42 observations in the 10-minute model). We chose the optimal constraint function based on the AIC criterion. In order to check the adequacy of functional constraints, we performed the heteroscedasticity and autocorrelation robust weight specification test (hARh) (Ghysels et al., 2016). If a model did not pass the test, we computed the 'unrestricted' MIDAS model, imposing no constraints on the regression parameters (see: Foroni et al., 2011) for the comparison of the unrestricted MIDAS models with the models with the Almond constraints). Next, the forecasts for the chosen model were generated. Our preliminary research demonstrated that the best results were obtained for the AR-MIDAS (not the simple MIDAS), therefore we used it. As the liquidity and volatility measures are non-stationary, we obtained the first differences in the variables. In the estimation of the AR-MIDAS models and the generation of the forecasts, we used the following R packages: midasr (Ghysels et al., 2016), forecast (Hyndman et al., 2019; Hyndman and Khandakar, 2008) and highfrequency (Boudt et al., 2018).

4.2. Vector autoregressive model

In the next step, we also computed forecasts of liquidity and volatility using the vector autoregression (further: the VAR). The VAR model has the following form:

$$\begin{cases} \Delta VOL_t = \sum_{i=1}^{k} \alpha_{1i} \Delta VOL_{t-i} + \sum_{i=1}^{k} \beta_{1i} \Delta LIQ_{t-i} + \epsilon_{1t} \\ \Delta LIQ_t = \sum_{i=1}^{k} \alpha_{2i} \Delta VOL_{t-i} + \sum_{i=1}^{k} \beta_{2i} \Delta LIQ_{t-i} + \epsilon_{1t} \end{cases}, \tag{9}$$

where ΔVOL_t denotes the change of volatility in day t, ΔLIQ_t is the change of liquidity between day $t-1$ and t, ϵ_t is the $iidN(0;\sigma)$ term, and $k \leq 5$. Two R packages were applied: vars (Pfaff, 2008a; 2008b) and VAR.etp (Kim, 2014).

Comparing the forecasts from the VAR with the forecasts from the MIDAS enabled us to check whether it was better to use only daily data on volatility and liquidity, or daily and intradaily data. The lag length was determined on the basis of the AIC criterion with the maximum allowed length of the lag being 5 days. Thus, we assumed that the impact of the information from a period longer than a week was not significant for the prediction purposes. In order to maintain consistency with previous approaches, the length of the out-of-sample interval was 100 days. We generated the one-day-ahead forecasts and computed the mean absolute error, MAE, and root mean square errors, RMSE (Hyndman & Koehler, 2006), according to the following formulas:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_{T+i} - \hat{y}_{T+i}|, \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{T+i} - \hat{y}_{T+i})^2}, \tag{11}$$

where y_{T+i} denotes the i-th value of the out-of-sample dependent variable, $\hat{y}_{T+i}$ its forecast, and N is the number of the out-of-sample length (in our case $N = 100$).

4.3. Autoregressive model and the naive approach

Eventually, we investigated whether the presence of regressors in the model improved the forecasts. To verify this, we compared the forecasts from the AR-MIDAS models with those obtained by means of the simple autoregressive models and the naive approach. By the ‘simple AR’ we mean the models with autoregressive variables in daily frequency only, without the MIDAS part. The lag number in the AR model was determined on the basis of the AIC information criterion, with the restriction on the maximum length of the lag to 5 days (as in the case of the VAR model). The length of the out-of-sample interval was 100 days. We generated the forecasts for one day ahead ($h = 1$), and computed the RMSE and the MAE errors.

5. Empirical results

In the empirical part of the paper, we estimated two sets of models – one in terms of the changes in volatility, and the other in terms of the changes in liquidity:

1. The change in the daily volatility approximated by RV is modelled as an autoregressive process with high-frequency explanatory variables involving changes in the 10-, 30- or 60-minute liquidity measured as the CQS:

$$\Delta RV_t = \alpha_i \Delta RV_{t-1} + \sum_{j=0}^{l} \beta_j \Delta LIQ_{m-j} + \epsilon_t. \tag{12}$$

   The alternative models are the VAR model with the CQS daily liquidity measure, a simple AR model for RV, and the naive approach.

2. The change in the daily liquidity represented by the CQS proxy was modelled as an autoregressive process with high-frequency explanatory variables standing for the changes in volatility, proxied by the squared returns calculated in 10-, 30- or 60-minute intervals (SQRET):

$$\Delta LIQ_t = \alpha_i \Delta LIQ_{t-1} + \sum_{j=0}^{l} \beta_j \Delta SQRET_{m-j} + \epsilon_t. \tag{13}$$

The following were the alternatives: the VAR model with RV calculated using data of the same frequency as the CQS, the AR models for the liquidity measure, and the naive approach.

We generated the forecasts from each model for each measure considered and we compared the forecast errors, calculated as the MAE and the RMSE. To assess the forecast ability of the models, in the first step we computed the ratio of the errors from the AR-MIDAS, the VAR and the AR to the naive errors. Additionally, we used the Diebold-Mariano test to verify the hypotheses that the values of the errors produced from the MIDAS model are smaller than the ones produced by the VAR, AR and the naive methods.

5.1. Forecasting ability of volatility models

The results of the forecasting study for realised volatility are presented in Table 1. The table shows the relationship between the one-step-ahead forecast errors from the AR-MIDAS model and the alternative approaches, i.e. the VAR model, the AR model and the naive approach. We considered forecasts of the volatility estimates, i.e. the daily RV, calculated on the basis of the 60-, 30- or 10-minute data. The sampling frequencies of the variables used in the models were consistent: the daily RV was based on the same frequency as the explanatory variables used in the models. With regard to the VAR models and all the sampling frequencies, the out-of-sample relative forecast errors were larger than 1. It means that in all the cases, the forecasts obtained by means of the AR-MIDAS models were less accurate than those received from the VAR models. The higher the sampling frequency, the larger the discrepancy in the predictions observed in the case of the VAR.

Table 1. Forecasting ability of the AR-MIDAS method compared to the VAR, AR and naive methods – realised volatility

Comparison method	Frequency of regressor	Alternative model		
		VAR	AR	NAIVE
RMSE	60 min	113.74%	114.84%	55.89%
	30 min	116.42%	117.31%	57.27%
	10 min	123.01%	123.70%	61.58%
MAE	60 min	110.78%	112.90%	58.25%
	30 min	114.63%	116.58%	60.20%
	10 min	120.93%	122.26%	63.85%
Diebold-Mariano test	60 min	1.69%	0.00%	98.31%
	30 min	0.85%	0.00%	95.76%
	10 min	0.00%	0.00%	94.92%

Note. The upper and middle parts of the table present the values of the relative errors of the prognosis calculated as a percentage ratio of the one-ahead forecast error from the AR-MIDAS model for volatility and from one of the alternative models: 1) the VAR model 2) the AR model or 3) the naive approach to the realised volatility value. Two types of errors are provided, i.e. the RMSE and MAE errors. The lower part of the table presents the percentage of cases where the forecast errors from the AR-MIDAS model were more accurate than the forecasts from the VAR model, the AR model or the naive forecasts. This comparison was performed for the MAE error on the basis of the Diebold-Mariano test. The numbers represent the percentage of cases where the AR-MIDAS model proved more effective than any of the alternatives. The results are shown separately for different sampling frequencies: 60-, 30- and 10-minute frequencies.

Source: authors' calculations.

When we compared the AR-MIDAS with a simple AR model without the MIDAS part, the results were similar. The values of forecasting errors from the AR-MIDAS models were definitely higher for the 10-minute data. The forecasts of daily RV-generated on the basis of the AR-MIDAS model were generally less accurate than those based on the simple AR model.

The results were quite opposite, however, when we compared the AR-MIDAS to the naive approach. In all the cases, the relative forecast errors were less accurate in the case of the AR-MIDAS than the naive approach. The forecasts based on the AR-MIDAS model were more accurate for shorter forecast horizons. Also, an improvement was observed in the forecast accuracy when the frequency of the explanatory variable was lower (e.g. 60-minute frequencies were used instead of 10-minute ones). Thus, the AR-MIDAS model proved to have an advantage over the naive method which, on the other hand, diminished as the VAR or the AR specification was used.

We also considered a different forecast error measure, i.e. the MAE, and examined the robustness of the results (see the middle part of Table 1). The results demonstrated that the VAR model allowed the generation of more accurate forecasts than the AR-MIDAS model. Moreover, no changes were observed in the results for the AR model nor the naive approach. The forecasting ability of the former is always higher than that of the AR-MIDAS specification, while the opposite is true for the latter.

We also applied the Diebold-Mariano test (Diebold, 2015) to compare the predictive accuracy and to verify whether the differences in the forecast errors resulting from the AR-MIDAS and those resulting from the three remaining approaches were significantly different from 0. We used a one-sided test where the null hypothesis stated that there were no differences between the two series of forecast errors, while the alternative hypothesis stated that the forecast errors of the AR-MIDAS model were less significant than those of the VAR model, AR model or the naive method. The test was applied to the forecasts generated separately for each stock and the results were averaged across the sample. The final result showed how often the predictive accuracy of the AR-MIDAS model was higher than that obtained from the remaining models in the cross section.

The lower part of Table 1 shows the results of the Diebold-Mariano test. We found that the AR-MIDAS model is more accurate only when compared with the naive approach. A simple AR model is always more accurate than the AR-MIDAS model, while the VAR model for the same horizon is almost always more accurate than the AR-MIDAS model. We also argue that, based on the results of the Diebold-Mariano test for errors, there is no need to employ a mixed-data sampling model in this particular forecasting case. The sole application of a simple AR or VAR model would generate more accurate forecasts of volatility.

5.2. Forecasting ability of liquidity models

In this section of the paper we consider forecasts of liquidity. The upper part of Table 2 shows the relative forecasting RMSE. We found that in terms of liquidity forecasts, the VAR and AR models were always more accurate than the AR-MIDAS model. Similarly to the volatility forecasting, the naive approach generated less accurate liquidity forecasts than those obtained on the basis of the AR-MIDAS model in all the considered frequencies.

Additionally, as in the case of the volatility forecasting, we investigated the relative MAE errors (see the middle part of Table 2). Here the results were slightly different:

in most cases the errors resulting from the application of the VAR model were greater than those resulting from the use of the AR-MIDAS model. The only exceptions were the forecasts for one day in 10-minute frequencies The same results were obtained for simple AR models, where the relative errors were lower than 1%, which indicated a slight predominance of the AR-MIDAS model. As far as the RMSE errors were concerned, the naive approach was still less accurate than the AR-MIDAS model.

We also provide the results of the Diebold-Mariano test. The lower part of Table 2 presents the percentage of cases where forecasts obtained by means of the AR-MIDAS models were of higher accuracy than the forecasts obtained by means of the alternative models. We found that as regards both the VAR and the AR models, in most cases their accuracy was higher than that of the AR-MIDAS. When the naive model was considered, its accuracy was in all cases lower than that of the AR-MIDAS.

Table 2. The forecasting ability of the AR-MIDAS compared to the VAR, AR and naive methods: liquidity

Comparison method	Frequency of regressor	Alternative model		
		VAR	AR	NAIVE
RMSE	60 min	105.49%	105.62%	58.24%
	30 min	106.06%	106.16%	57.23%
	10 min	107.89%	107.98%	58.38%
MAE	60 min	96.22%	96.43%	57.26%
	30 min	99.13%	99.24%	56.08%
	10 min	100.42%	100.59%	56.54%
Diebold-Mariano test	60 min	22.88%	21.19%	100.00%
	30 min	11.86%	11.02%	100.00%
	10 min	6.78%	7.63%	100.00%

Note. The upper and middle parts of the table present the values of the relative errors calculated as a percentage ratio of the one-ahead forecast error from the AR-MIDAS model for liquidity and from one of the alternative models, namely the VAR model, the AR model or the naive approach to the realised liquidity value. Two types of errors are provided, i.e. RMSE and MAE errors. The lower part of the table presents the percentage of cases where forecast errors from the AR-MIDAS were of higher accuracy than the forecasts from the VAR model, the AR model or the naive forecasts. This comparison was performed for the MAE error on the basis of the Diebold-Mariano test. The numbers represent the percentage of cases where the AR-MIDAS model was more effective than any of the alternatives. The results are shown separately for 60-, 30- and 10-minute frequencies.

Source: authors' calculations.

It is also worth noting that the percentage of cases where the AR-MIDAS outperformed the VAR or the AR models was higher when liquidity was predicted using intradaily volatility rather than vice versa. This indicates that the changes in intradaily volatility influence the dynamics of daily liquidity more often than the changes of intradaily liquidity influence daily volatility. This suggests that investors observe the changes of prices during the day and on this basis make decisions as to whether to change their position in the asset. In other words – what influences the decision to change the position is more often the movement of prices rather than the interest of other market participants.

5.3. Are liquidity and volatility self-explanatory processes? A comparison of the AR and the VAR models

Research carried out to date shows that for both volatility and liquidity forecasting, the VAR and the AR models generate on average more accurate forecasts than the AR-MIDAS models. On the other hand, the latter are better in terms of forecast accuracy than forecasts generated by means of the naive approach. However, the question as to which out of the two, the VAR or the AR, is more effective in forecasting either volatility or liquidity, remains to be answered.

The 'Volatility' column of Table 3 presents a comparison of the forecast errors, i.e. the relative forecast errors from the volatility forecasts based on VAR and AR models. We found that, regardless of the frequency of the data and the forecast error measure, AR models generate a slightly lower number of errors.

The same approach was applied to liquidity forecasts. The results are presented in the 'Liquidity' column of Table 3. They show that all the fractions are very close to 1%, which means that there is no significant difference between forecasts generated through the VAR or the AR model.

Table 3. Forecast error of volatility and liquidity changes: comparison of the VAR and AR model

Error type	Data frequency	Volatility	Liquidity
RMSE	10 min	100.56%	100.08%
	30 min	100.79%	100.07%
	60 min	101.16%	100.11%
MAE	10 min	101.06%	100.16%
	30 min	101.81%	100.10%
	60 min	102.42%	100.21%

Note. The table presents the percentage ratio of the RMSE and MAE forecast errors from the VAR and AR models for different forecast horizons. In the VAR model we take into account the lagged daily volatility (RV) and liquidity (CQS).

Source: authors' calculations.

The above leads to the conclusion that in the case of volatility, the AR model might generate slightly better forecasts in terms of accuracy than the VAR model, while in the case of liquidity, the forecasts from both models are equally accurate.

6. Discussion and conclusions

Literature on the undertaken subject provides evidence for one-direction or bi-directional causality between volatility and liquidity. The research presented in this paper aimed to verify whether this dependence could be used to improve forecasts of both volatility and liquidity. Four approaches were considered: the first was based on a mixed sampling of data where daily forecasts of volatility (or liquidity) were generated on the basis of the intraday liquidity (or volatility) measures. The second approach was a VAR model based on daily variables only. The two alternatives – a simple AR model and the naive approach – employed only previous realisations of the processes for which the forecasts were generated.

We have found that although using cross-dependency between volatility and liquidity has its advantages, the employment of mixed-data sampling models is not justified. MIDAS models provide more accurate forecasts than those based on the naive approach. However, in the case of volatility forecasts, both the VAR models with lagged volatility and liquidity and the AR models with lagged liquidity generate errors of lower values than the forecasts based on the MIDAS specifications. With regard to liquidity forecasts, there is no significant difference in forecast accuracy between the MIDAS models and the VAR or AR specification. Thus, the values of forecast errors of volatility are lower when one uses previous values of volatility and previous liquidity data in daily frequency only. However, the computational burden and the associated effort of employing the MIDAS model is much greater than that entailed by the simple AR or VAR model. When only the two latter are compared, the ratio of their respective errors is close to one, indicating that there are only negligible differences between both approaches.

Additionally, the prevalence of the VAR and AR models over the MIDAS model becomes even more evident with the application of higher-frequency data (e.g. 10-minute instead of 60-minute data). Results thus produced are important for investors as well as risk managers who might be wondering if it is worth employing more advanced models that require enormous computing power. Our empirical study shows that in the case of liquidity and volatility forecasting, the gains obtained from the use of MIDAS models are rather negligible. The outcome also sheds some light on the behavioural aspect of investing on the WSE. Considering the fact that the percentage of cases where the AR-MIDAS outperformed the VAR or AR models was higher when liquidity was predicted using intradaily volatility than when volatility was predicted using intradaily liquidity, the conclusion is that what influences the decision on the change of the position is more often the movement of prices rather than the interest of other market participants. The authors' further research in this area will be devoted to examining the stability of these results by means of other volatility and liquidity measures.

Acknowledgements

This work was supported by the National Science Centre in Poland under grant no. UMO-2017/25/B/HS4/01546. We would like to thank the seminar participants at Adam Mickiewicz University (SEFIN), as well as the attendants of the INFINITI Conference in 2019 in Glasgow and the 26th International Conference of Forecasting Financial Markets in Venice, for their valuable comments. All errors are our own.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon request.

References

Ahn, H.-J., Cai, J., & Yang, C.-W. (2018). Which Liquidity Proxy Measures Liquidity Best in Emerging Markets?. *Economies*, *6*(4), 1–29. https://doi.org/10.3390/economies6040067.

Andersen, T. G., & Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review*, *39*(4), 885–905. https://doi.org/10.2307/2527343.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., & Diebold, F. X. (2006). Volatility and Correlation Forecasting. In G. Elliott, C. W. J. Granger, A. Timmermann (Eds.), *Handbook of Economic Forecasting* (vol. 1, pp. 777–878). Elsevier. https://doi.org/10.1016/S1574-0706(05)01015-3.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility. *Review of Economics and Statistics*, *89*(4), 701–720. https://doi.org/10.1162/rest.89.4.701.

Andreou, E., Ghysels, E., & Kourtellos, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics*, *158*(2), 246–261. https://doi.org/10.1016/j.jeconom.2010.01.004.

Asgharian, H., Hou, A. J., & Javed, F. (2013). The importance of the macroeconomic variables in forecasting stock return variance: A GARCH-MIDAS approach. *Journal of Forecasting*, *32*(7), 600–612. https://doi.org/10.1002/for.2256.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2009). Realized kernels in practice: trades and quotes. *Econometrics Journal*, *12*(3), C1–C32. https://doi.org/10.1111/j.1368-423X.2008.00275.x.

Będowska-Sójka, B., & Kliber, A. (2019). The Causality between Liquidity and Volatility in the Polish Stock Market. *Finance Research Letters*, *30*, 110–115. https://doi.org/10.1016/j.frl.2019.04.008.

Boudt, K., Cornelissen, J., Payseur, S., Nguyen, G., & Schermer, M. (2018). *Highfrequency: Tools for Highfrequency Data Analysis. R Package version 0.5.3*.

Breitung, J., & Roling, C. (2015). Forecasting inflation rates using daily data: A nonparametric MIDAS approach. *Journal of Forecasting*, *34*(7), 588–603. https://doi.org/10.1002/for.2361.

Brunnermeier, M. K., & Pedersen, L. H. (2009). Market Liquidity and Funding Liquidity. *Review of Financial Studies*, *22*(6), 2201–2238. https://doi.org/10.1093/rfs/hhn098.

Chordia, T., Subrahmanyam, A., & Anshuman, V. R. (2001). Trading activity and expected stock returns. *Journal of Financial Economics*, *59*(1), 3–32. https://doi.org/10.1016/S0304-405X(00)00080-5.

Chung, K. H., & Zhang, H. (2014). A simple approximation of intraday spreads using daily data. *Journal of Financial Markets*, *17*, 94–120. https://doi.org/10.1016/j.finmar.2013.02.004.

Diebold, F. X. (2015). Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests. *Journal of Business & Economic Statistics*, *33*(1), 1–24. https://doi.org/10.1080/07350015.2014.983236.

Díaz, A., & Escribano, A. (2020). Measuring the multi-faceted dimension of liquidity in financial markets: A literature review. *Research in International Business and Finance*, *51*. https://doi.org/10.1016/j.ribaf.2019.101079.

Engle, R. F., Ghysels, E., & Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *The Review of Economics and Statistics*, *95*(3), 776–797.

Faff, R. W., Hillier, D., & Hillier, J. (2000). Time Varying Beta Risk: An Analysis of Alternative Modelling Techniques. *Journal of Business Finance & Accounting*, *27*(5–6), 523–554. https://doi.org/10.1111/1468-5957.00324.

Ferrara, L., & Marsilli, C. (2013). Financial variables as leading indicators of GDP growth: Evidence from a MIDAS approach during the great recession. *Applied Economics Letters*, *20*(3), 233–237. https://doi.org/10.1080/13504851.2012.689099.

Fong, K. Y. L., Holden, C. W., & Trzcinka, C. A. (2017). What are the best liquidity proxies for global research? *Review of Finance*, *21*(4), 1355–1401. https://doi.org/10.1093/rof/rfx003.

Foroni, C., Marcellino, M., & Schumacher, C. (2011). *U-MIDAS: MIDAS regressions with unrestricted lag polynomials* (Discussion Paper Series 1: Economic Studies, No 35). https://www.econstor.eu/bitstream/10419/55529/1/685618153.pdf.

Fuertes, A.-M., & Olmo, J. (2012). Exploiting Intraday and Overnight Price Variation for Daily VaR Prediction. *Frontiers in Finance and Economics*, *9*(2), 1–31.

Ghysels, E., Kvedaras, V., & Zemlys, V. (2016). Mixed frequency data sampling regression models: The R package midasr. *Journal of Statistical Software*, *72*(4), 1–35. https://doi.org/10.18637/jss.v072.i04.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004). *The MIDAS Touch: Mixed Data Sampling Regression Models* (CIRANO Working Papers, 2004s-20). https://www.cirano.qc.ca/files/publications/2004s-20.pdf.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, *131*(1–2), 59–95. https://doi.org/10.1016/j.jeconom.2005.01.004.

Ghysels, E., Sinko, A., & Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, *26*(1), 53–90. https://doi.org/10.1080/07474930600972467.

Giot, P. (2005). Market risk models for intraday data. *European Journal of Finance*, *11*(4), 309–324. https://doi.org/10.1080/1351847032000143396.

Gold, N., Wang, Q., Cao, M., & Huang, H. (2017). Liquidity and volatility commonality in the Canadian stock market. *Mathematics-in-Industry Case Studies*, *8*(7), 1–20. https://doi.org/10.1186/s40929-017-0016-9.

Hautsch, N., & Jeleskovic, V. (2008). *Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models* (SFB 649 Discussion Papers). http://dx.doi.org/10.2139/ssrn.1292493.

Hiemstra, C., & Jones, J. D. (1994). Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation. *Journal of Finance*, *49*(5), 1639–1664. https://doi.org/10.2307/2329266.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., & Yasmeen, F. (2019). *Forecast: Forecasting functions for time series and linear models. R package version 8.5*.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, *27*(3), 1–22. https://doi.org/10.18637/jss.v027.i03.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001.

Kim, J. H. (2014). *VAR.etp: VAR modelling: estimation, testing, and prediction. R package version 0.7*.

Kim, H. H., & Swanson, N. R. (2018). Methods for backcasting, nowcasting and forecasting using factor-MIDAS: With an application to Korean GDP. *Journal of Forecasting*, *37*(3), 281–302. https://doi.org/10.1002/for.2499.

Laurent, S., & Violante, F. (2012). Volatility forecasts evaluation and comparison. *WIREs Computational Statistics*, *4*(1), 1–12. https://doi.org/10.1002/wics.190.

Liang, S. X., & Wei, J. K. C. (2012). Liquidity risk and stock returns around the world. *Journal of Banking and Finance*, *36*(12), 3274–3288. https://doi.org/10.1016/j.jbankfin.2012.07.021.

Ma, R., Anderson, H. D., & Marshall, B. R. (2018). Stock market liquidity and trading activity: Is China different? *International Review of Financial Analysis*, *56*, 32–51. https://doi.org/10.1016/j.irfa.2017.12.010.

Ma, Y.-r., Ji, Q., & Pan, J. (2019). Oil financialization and volatility forecast: Evidence from multidimensional predictors. *Journal of Forecasting*, *38*(6), 564–581. https://doi.org/10.1002/for.2577.

Maas, B. (2019). Short-term forecasting of the US unemployment rate. *Journal of Forecasting*, *39*(3), 394–411. https://doi.org/10.1002/for.2630.

Monteforte, L., & Moretti, G. (2013). Real-time forecasts of inflation: The role of financial variables. *Journal of Forecasting*, *32*(1), 51–61. https://doi.org/10.1002/for.1250.

Ong, M. A. (2015). An information theoretic analysis of stock returns, volatility and trading volumes. *Applied Economics*, *47*(36), 3891–3906. https://doi.org/10.1080/00036846.2015.1019040.

Pástor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111*(3), 642–685. https://doi.org/10.1086/374184.

Pfaff, B. (2008a). *Analysis of Integrated and Cointegrated Time Series with R* (2nd edition). Springer.

Pfaff, B. (2008b). VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software*, *27*(4), 1–32. https://doi.org/10.18637/jss.v027.i04.

Rösch, C. G., & Kaserer, C. (2014). Market liquidity in the financial crisis: The role of liquidity commonality and flight-to-quality. *Journal of Banking and Finance*, *37*(7), 2284–2302. https://doi.org/10.1016/j.jbankfin.2013.01.009.

Santos, D. G., & Ziegelmann, F. A. (2014). Volatility forecasting via MIDAS, HAR and their combination: An empirical comparative study for IBOVESPA. *Journal of Forecasting*, *33*(4), 284–299. https://doi.org/10.1002/for.2287.

Smith, P. (2016). Google's MIDAS touch: Predicting UK unemployment with Internet Search Data. *Journal of Forecasting*, *35*(3), 263–284. https://doi.org/10.1002/for.2391.

Tsui, A. K., Xu, C. Y., & Zhang, Z. (2018). Macroeconomic forecasting with mixed data sampling frequencies: Evidence from a small open economy. *Journal of Forecasting*, *37*(6), 666–675. https://doi.org/10.1002/for.2528.

Yeyati, E. L., Telia, D., & Schmukler, S. L. (2008). Emerging Market Liquidity and Crises. *Journal of the European Economic Association*, *6*(2–3), 668–682.

Przegląd Statystyczny. Statistical Review, 2022, vol. 69, 2, 20–36 DOI: 10.5604/01.3001.0016.0364
Received: 22.08.2022, revised: 28.09.2022, accepted: 1.10.2022

On planning production and distribution with disrupted supply chains

Przemysław Szufel[a]

Abstract. This paper presents a model for short-term time-horizon production and distribution planning of a manufacturing company located in the middle of a supply chain. The model focuses on an unbalanced market with broken supply chains. This reflects the state of the current post-COVID-19 economy, which is additionally struggling with even more uncertainty and disruptions due to the Russian aggression against Ukraine. The manufacturer, operating on the post-pandemic and post-war market, on the one hand observes a soaring demand for its products, and on the other faces uncertainty regarding the availability of components (parts) used in the manufacturing process. The goal of the company is to maximise profits despite the uncertain availability of intermediate products. In the short term, the company cannot simply raise prices, as it is bound by long-term contracts with its business partners. The company also has to maintain a good relationship with its customers, i.e. businesses further in the supply chain, by proportionally dividing its insufficient production and trying to match production planning with the observed demand. The post-COVID-19 production-planning problem has been addressed with a robust mixed integer optimisation model along with a dedicated heuristic, which makes it possible to find approximate solutions in a large-scale real-world setting.

Keywords: production, optimisation techniques, simulation modelling, programming models, transportation economics

JEL: C44, C61, L90

1. Introduction

The COVID-19 pandemic has changed the way markets and economies function across almost all industry branches. Severe disturbances in how supply chains operate can be currently observed all over the world. The just-in-time supply model is no longer feasible for many companies (Brakman et al., 2020). This complicated situation even worsened due to the Russian aggression against Ukraine, which caused damages to Ukraine's economy, the seizure of export routes and further complications to the processes of production and distribution of goods, the latter being predominantly the consequence of global sanctions against the aggressor (Mbah & Wasum, 2022).

Another phenomenon observed across many industry sectors is decreased availability of consumer products such as electronics, computers, or vehicles

[a] SGH Warsaw School of Economics, Collegium of Economic Analysis, al. Niepodległości 162, 02-554 Warszawa, e-mail: pszufe@sgh.waw.pl, ORCID: https://orcid.org/0000-0001-9525-3497.

(Chen et al., 2021). The globally-observed changes in the availability of raw materials and semi-products of critical components have changed the way manufacturing companies should operate and intermediate products or final goods be provisioned to companies further in the supply chain, customers and resellers.

Several papers show that the COVID-19 pandemic, and later the war in Ukraine, have drastically influenced the efficiency of global supply chains. Cai and Luo (2020) reviewed the impact of COVID-19 on the manufacturing chain, pointing out its negative influence on raw materials, spare parts, intermediate products and workforce availability. The authors further noticed that on the one hand, companies had to adapt to the chain disturbances, and on the other needed new methods to enhance supply chain resilience. The final conclusion of their study is that in the post-crisis world, the manufacturing supply chain is likely to become regionalised and digitalised.

Singh et al. (2021) pointed out that the COVID-19 pandemic was continuously causing disturbances across all the levels of the economy. It affected the access to crucial resources (employees, logistics, raw materials), essential items (basic food commodities, perishable food items, medicines, diagnostic equipment, personal protection equipment), primary economic sectors (aviation, railway, agriculture, healthcare, FMCG) as well as sectors playing a significant role in modern economies (hospitality, construction, information technology, automotive and textile manufacturing). Similar conclusions could also be found in Queiroz et al. (2020).

In order to mitigate supply chain disturbances, Paul and Chowdhury (2021) considered a theoretical scenario of a manufacturing system where under normal circumstances the production was higher than the demand, but the situation reversed as a result of the COVID-19 pandemic, i.e. the demand surged while the production capacity decreased (supply disturbances). To solve this problem, the authors proposed a model where the reserve storage capacity was expanded and purchases of raw materials from more expensive suppliers were increased with the purpose of speeding up the production process. In their approach, the production plan was represented by a constrained nonlinear optimisation problem which they solved with gradient methods. A simplified version of this model was presented by Shahed et al. (2021), who based it on a profit-maximising manufacturer with a single supplier and a single retailer. This research pointed out the importance of re-implementing inventory management policies in manufacturing companies.

Tsolas and Hasan (2021) proposed economic survivability (understood as a point where a business ceases to be profitable) model to explain decisions of a company operating in a market with high fluctuations of raw materials availability and demand. The authors built a survivability-maximising optimisation model and showed that a company should balance the allocation of its manufacturing plants across multiple regions and ensure diversified supply chain connections between suppliers of raw materials and the factories – even if it leads to a decrease in the overall profit. On the other hand, Li et al. (2020) stressed the role of intelligent manufacturing as a proactive method to mitigate production disruptions caused by a pandemic. They proposed to implement a continuous decision-making model for determining the optimal deployment of resources to strengthen the existing industrial network.

The literature shows that manufacturing and distribution companies are currently facing two types of uncertainties: on the one hand, lack of raw material semi-products and intermediate and key components that affect the production activity, and on the other huge fluctuations in demand coupled with product shortages in many consumer markets that is disruptive to the distribution activities.

For the purpose of this paper, we selected a manufacturing company that experienced a soaring demand for its products and at the same time faced shortages of critical components, which made it impossible for it to operate at full production capacity. The goal of the company was to maximise profits despite uncertain availability of intermediate products. The research presented in the paper focuses on a short-term decision-making horizon. By 'short-term' we mean that the prices agreed upon by the company and its customers were fixed, i.e. the company was bound by long-term pricing contracts. For this reason, and despite limited supply, the product allocation problem could not be simply addressed by raising prices, as the company had to take into consideration long-term relationships with its customers. Moreover, huge fluctuations in the demand were observed on the market and the availability of components critical to the production could not be guaranteed. Businesses struggling with this kind of problems are, for example, automotive dealers, medicine producers and producers of electronic devices.

Our paper presents a novel approach, featuring a short-term model of a market with fixed demand, an insufficient supply of goods and a reduced price flexibility. This approach has been selected in order to analyse the decision-making process in the current post-pandemic economy that is additionally struggling with the effects of the Russian aggression in Ukraine. The manufacturer chosen for the purpose of the study was able to manufacture only a limited number of goods that had to be distributed among companies located further in the supply chain. The goal of the manufacturer was on the one hand to minimise the potential frustration of its customers and on the other to maximise profits.

In order to show our problem in a business setting, let us consider an automotive manufacturer that has long-term contracts and business relations with car dealers. In the short term, the manufacturer cannot adjust the price list. However, since cars yield different profit margins, they can be distributed in many ways among different dealers. The manufacturer cannot ignore that fact that each car dealer made some pre-orders or entered into long-term contracts when the economic situation was different. However, since there generally is a significant shortage of cars, we can observe that several producers e.g. started manufacturing vehicles with different equipment than originally planned (Boston, 2022). For instance, they provide customers with vehicles that have different engine types or are of different colours than it was stated in the original order.

The goal of the paper is to propose an optimisation approach to address the problem of a manufacturer experiencing disruptions of supplies and at the same time soaring demand, all in a short-term decision-making horizon. In order to address the uncertainty of critical component availability, the study adopted the robust optimisation approach of Beyer and Sendhoff (2007). The model assumes that manufacturers of critical components might adopt a similar strategy as the company selected for the purposes of the study, i.e. provide the manufacturer of end products with slightly different components than originally requested.

The paper is constructed in the following way: Introduction is followed by Section 2, where a mathematical model formulation is proposed, Section 3 presents the results of numerical experiments, and Section 4 comprises the conclusions of the study.

2. Problem statement and model

As mentioned before, the profile of the company analysed in this study is one that uses several intermediate components (parts) to manufacture a single product. Companies meeting this criterion include manufactures of computing hardware, cars, e-scooters, electronic appliances and furniture. Our model makes allowances for the fact that manufactures of this kind usually have broad, long-established dealership networks, in the framework of which business relationships have often lasted for a long time and which are an important part of these companies' values. As a consequence of the fact that there are shortages of goods in the market, the customers of such manufacturers are usually willing to accept end products that are slightly different than the original order.

2.1. Managing baseline demand

We are considering a manufacturing company with demand $v_{dn} \in \mathbb{N}_0$, where $d \in D$ is a customer from customer base D, and $n \in N$ is a product from product set N (we use $\mathbb{N}_0$ to denote non-negative integers). The demand of customer $d \in D$ for product $n \in N$ is denoted as v_{dn}, and hence the overall demand is represented by matrix $V = [v_{dn}]$.

In order to manufacture the goods, a set of critical components K is required. The technology matrix is represented by $A = [a_{kn}]$, where $[a_{kn}]$ stands for the number of parts of type k to manufacture good n. $b_k \in \mathbb{N}_0$ is the expected (optimistic) level of the critical component availability. The number of available parts is known only approximately due to disturbances on the market, and so the unknown perturbance is represented by $\xi_k \in \mathbb{N}_0$, $\xi_k \leq \psi_k$ with the maximum perturbance limit of $\psi_k \in \mathbb{N}_0$. Moreover, since there is a possibility of replacing some components with others, we assumed that there is a maximum total perturbation level $\Gamma \in \mathbb{N}_0$, such that $\sum_{k \in K} \xi_k \leq \Gamma$. Hence, for the considered availability of components b, we define the following uncertainty set $\mathbf{U}$ known in the literature (e.g. Li et al., 2011) as the boxed-polyhedral uncertainty:

$$\mathbf{U} = \left\{ \boldsymbol{\xi} \,\middle|\, \mathbf{0} \leq \boldsymbol{\xi} \leq \boldsymbol{\psi} \wedge \left|\left|\boldsymbol{\xi}\right|\right|_1 \leq \Gamma \right\}, \quad (1)$$

where we use bold font to represent the vectors of values, i.e. $\boldsymbol{\psi} = [\psi_k]$ and $\boldsymbol{\xi} = [\xi_k]$. Moreover, the $||\,.\,||_1$ notation denotes L^1 norm, i.e. $\left|\left|\boldsymbol{\xi}\right|\right|_1 = \sum_{k \in K} \xi_k$.

The studied manufacturer, as already mentioned, operates on a market with significant product shortages, and therefore has to fulfill the demand by offering similar but slightly different products that will be further called 'substitutes'. The product substitution matrix is denoted by $S = [s_{ij}]$, where $s_{ij} = 1$ means that product $i \in N$ can be replaced by product $j \in N$, and 0 means that no replacement is possible. Please note that a product can always be a replacement for itself, and hence the S matrix has 1's on the diagonal (i.e. $s_{ii} = 1$ for $i \in N$). Additionally, we assume that substitutability is a symmetric relation (i.e. $s_{ij} = s_{ji}$), and hence S is symmetric. The company needs to decide how many goods to manufacture in order to satisfy the demand to the fullest possible extent. The production volume is represented by $\mathbf{y} = [y_n]$, $y_n \in \mathbb{N}_0$, and the allocation of those products to customers by $X = [x_{dn}]$, $x_{dn} \in \mathbb{N}_0$. Obviously, the demand can only be fulfilled to the extent allowed by the volume of production ($\sum_{d \in D} x_{dn} \leq y_n$). However, since we are considering a market that is unbalanced in the short term, the demand does not need to be fulfilled with the products that have been actually ordered, as substitutes can be used instead (for example, a green vehicle can be offered to the customer instead of a blue one). Each product $n \in N$ has a unit cost of manufacturing $c_n \geq 0$, and can be sold to customer $d \in D$ at price $p_{dn} \geq c_n$. The goal of the decision-maker, as already stated, is twofold: to maximise the profit and to maintain good long-term relationship with the customer network. For this reason, the business problem is a two-criteria optimisation, but in our approach, the second criterion (the degree to which customers' demand is met) is modelled by a constraint in the optimisation problem. The complete list of symbols used to describe the decision situation is presented in Table 1. Since we are considering manufacturing industries which have both large product portfolios (e.g. offer similar products in different colours or with slightly different technical specifications) and significant number of customers, the decision variables in the model are discrete rather than continuous.

2.2. Optimisation model

As mentioned before, the goal of decision-makers is twofold: besides maximising profits from sales, they strive to maintain good relations with their customer networks (by guaranteeing a minimal level of supply).

This can be presented as the following model:

$$\max \sum_{n\in N, d\in D} x_{dn}p_{dn} - \mathbf{c}^T\mathbf{y}, \tag{2}$$

subject to:

$$A\mathbf{y} \leq \mathbf{b} - \boldsymbol{\xi}, \tag{3a}$$

$$XS \leq VS, \tag{3b}$$

$$X \geq \gamma X^*, \tag{3c}$$

$$\sum_{d\in D} x_{dn} \leq y_n \ \forall n \in N, \tag{3d}$$

$$x_{dn} \in \mathbb{N}_0 \ \forall d \in D, n \in N, \tag{3e}$$

$$y_n \in \mathbb{N}_0 \ \forall n \in N, \tag{3f}$$

where $\boldsymbol{\xi}$ represents the boxed polyhedral uncertainty defined in Equation (1).

Function (2) represents profit maximisation. Typically, each customer $d \in D$ has a long-term relationship with the manufacturer, having negotiated an individual price p_{dn} for a particular product $n \in N$. This means that the price level can vary across the customer base. The profit is denoted as $\rho(X)$. Please note that the increase in the value of y without a corresponding drop in x will always lead to a decrease of the goal Function (2). Therefore, at optimal solution $(X^{opt}, \mathbf{y}^{opt})$, no unsold products will be manufactured, hence the following equation holds true:

$$\rho(X) = \sum_{n\in N, d\in D} x_{dn} p_{dn} - \mathbf{c}^T \mathbf{y} = \sum_{n\in N, d\in D} x_{dn}(p_{dn} - c_n). \tag{4}$$

Equation (3a) assumes that the production should not be greater than the uncertain availability of critical components, where $\mathbf{b}$ is the 'optimistic' availability of components, and $\boldsymbol{\xi}$ is the unknown perturbation. Equation (3b) defines the ranges for product substitution (which involves providing the customers with alternative products to the ones originally requested). Please note that the dimension of XS is $|D| \cdot |N|$ and, hence, in a given matrix row each value controls the total number of products within a group of substitutes for the product corresponding to the column. While Equation (3b) allows the free movement of products across customers, in practice there is still some minimal guaranteed level of matching the actual demand – here presented as γX^* in Equation (3c). Parameter $\gamma \in [0,1]$ represents the substitution rigidity for the minimal required allocation. $\gamma = 0$ means that customers can be freely offered substitutes, while $\gamma = 1$ says that the level of substitute products is minimal. X^* is the maximum feasible solution to the 'pessimistic' version of the problem (i.e. when $\boldsymbol{\xi} = \boldsymbol{\psi}$). One of the possible ways of calculating X^* will be shown in the subsequent part of the text. Finally, Equation (3d) ensures that the number of allocated stocks does not exceed the volume of the manufactured output.

Let us discuss a sample procedure for finding a feasible value of X^*. Since the size of X would be very large in practical applications, we propose a two-step procedure.

In the first step, a model is constructed that minimises the deviation of production from the current demand assuming pessimistic availability of the critical parts:

$$\min \sum_{n\in N} \left(\sum_{d\in D} v_{dn} - y_n\right)^2, \tag{5}$$

subject to:

$$A\mathbf{y} \le \mathbf{b} - \boldsymbol{\psi}, \tag{6a}$$

$$y_n \in \mathbb{N}_0. \tag{6b}$$

This model yields a pessimistic feasible value of production that matches demand V. We will denote that value as y^*. When the pessimistic value of production $\mathbf{y}^*$ is known, feasible allocation X^* can be calculated so that for each $n \in N$, the percentage deviation from the reported demand is minimised:

$$\min \sum_{d \in D; v_{dn} > 0} \left(\frac{v_{dn} - x_{dn}}{v_{dn}} \right)^2, \tag{7}$$

subject to:

$$x_{dn} \leq v_{dn}, \tag{8a}$$

$$\sum_{\mathrm{d} \in \mathrm{D}} x_{dn} \leq y_n^*, \tag{8b}$$

$$x_{dn} \in \mathbb{N}_0 \; \forall d \in D, n \in N. \tag{8c}$$

Table 1. Notation summary

Input variables	
$n \in N$	product type, where N is the set of considered products
$d \in D$	buyer, where D is the set of buyers
$k \in K$	part type (critical component) required to manufacture a given type of product, where K is the set of part types
$b_k \in \mathbb{N}_0$	optimistic-assumption availability of critical components k
$\xi_k \in \mathbb{N}_0$	unknown perturbation to the availability of parts; the perturbation vector is denoted as $\boldsymbol{\xi} = [\xi_k]$
$\psi_k \in \mathbb{N}_0$	maximum possible perturbation of the availability of components $\psi_k \geq 0$; maximum perturbation vector is denoted as $\boldsymbol{\psi} = [\psi_k]$
$\Gamma \in \mathbb{N}_0$	maximum L^1 norm of the possible perturbations $\lVert \boldsymbol{\xi} \rVert_1 \leq \Gamma$
$a_{kn} \in \mathbb{N}_0$	amount of critical parts k required to manufacture one unit of product n; the technology matrix is represented by $A = [a_{kn}]$
$s_{ij} \in \{0,1\}$	product substitution equivalent, where $s_{ij} = 1$ means that product i can be replaced by product j, where $i, j \in N$, and $s_{ij} = 0$ means that no replacement is possible, and the substitution matrix is represented by $S = [s_{ij}]$
$v_{dn} \in \mathbb{N}_0$	buyer demand d for product n; the demand matrix is represented by $V = [v_{dn}]$
p_{dn}	price acquired from selling product n to buyer d (prices vary across buyers due to different contract terms)
c_n	unit costs of manufacturing one unit of product n, in vector notation denoted as $\mathbf{c} = [c_n]$
x^*_{dn}	pessimistic level of the allocation of products; it can be represented by matrix $X^* = [x^*_{dn}]$
$\gamma \in [0,1]$	substitution rigidity; $\gamma = 0$ means that all customer orders can be replaced with substitutes, $\gamma = 1$ means that the level of substitutes will be minimised, and at least X^* will be fulfilled
$\rho(X)$	profit from production allocation X; optimistic and pessimistic profits are denoted as ρ^+ and ρ^-, respectively
Optimisation model variables	
y_n	production level of item n; it can be represented by vector $\boldsymbol{y} = [y_n]$
x_{dn}	number of goods of type n allocated to buyer d; it can be represented by matrix $X = [x_{dn}]$

Source: author's work.

Goal Function (7) involves minimising the percentage deviations of supply and demand. Note that for each $n \in N$, the exact solution to (7) can be easily achieved in three steps: (1) the proportional scaling of v_{dn} values in such a way that their sum is equal to y_n, (2) rounding those values down to the nearest integer and, finally, (3) redistributing the remaining product units starting from customers with the smallest orders. In practice, the decision-maker might decide to use substitution rigidity parameter γ to downscale the value of X^* and, as a result, offer more aggressively substitutes to their customer base instead of the ordered products, thereby generating a greater profit from the unbalanced market situation.

The dependencies between the subsequent modelling steps are presented in Figure 1. Firstly, we calculate the pessimistic production volume by solving the model presented in Equation 5. Secondly, we calculate the pessimistic amount of goods that will be made available to customers (please note that we are considering a post-pandemic economy with a scarcity of goods). It is important to remember that this solution will be feasible regardless of the observed perturbation value $\boldsymbol{\xi}$. Finally, since there is no more stable production capacity matching the demand, the company will manufacture product substitutes that will also be accepted by the market. This process is controlled by substitution rigidity parameter γ– its actual value will depend on the business objectives of the company.

It is assumed that since the demand on the market is high, customers will purchase substitute goods as long as Equation (3b) holds. Please note that due to uncertainty **U** (see Equation (1)), the decision-maker faces risk related to cash flow management and needs to adjust the production plan accordingly.

Given uncertainty set **U**, the pessimistic value of profits ρ^- can be calculated by solving the following model:

$$\rho^- = \min_{\xi \in U}\left\{max \sum\nolimits_{n \in N, d \in D} x_{dn} p_{dn} - \mathbf{c}^T \mathbf{y}; \quad subject\ to\ constr.\, 3a\text{–}\, 3f\right\}. \quad (9)$$

On the other hand, the optimistic value of profits ρ^+ can be computed by means of solving Equation (2), assuming that $\boldsymbol{\xi} = \mathbf{0}$. Bertsimas et al. (2016) point out that there are several approaches to reformulating a robust MILP optimisation model into a set of MILP models, but they all assume that perturbations are defined individually for each constraint (see e.g. Ben-Tal et al., 2009 or Li et al., 2011). Since the model presented in Equation (9) is the mixed integer programming, and the uncertainty set is defined across all constraints, it can only be solved through iterating solutions over the entire set **U**.

For larger sizes of uncertainty set **U**, iterating over all of its values is prohibitively computationally expensive. However, nearly optimal solution can be found by using a cutting plane heuristic similar to the one proposed by Bertsimas et al. (2016).

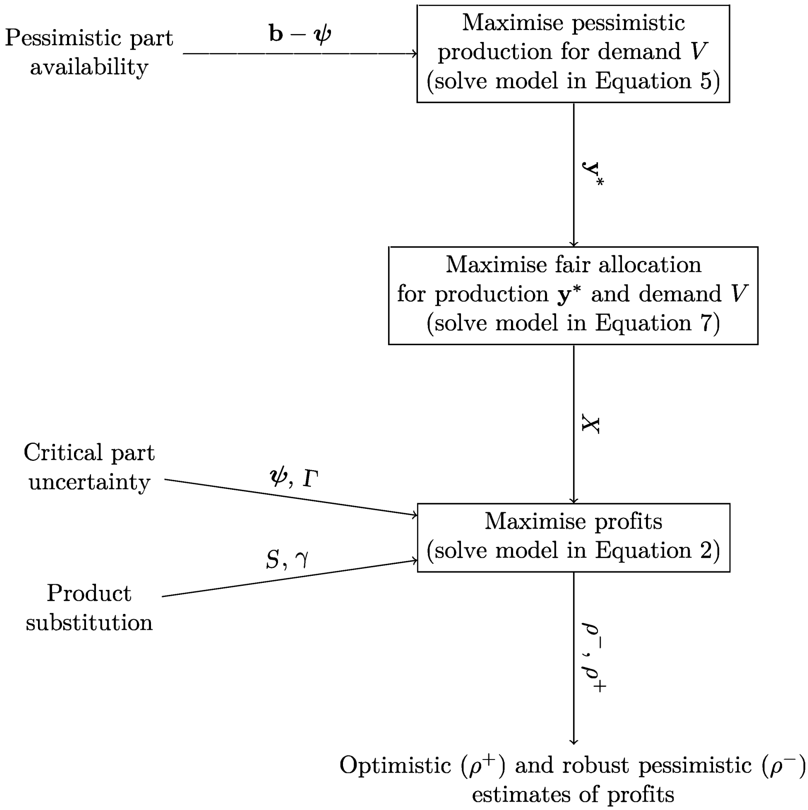
In this paper, the following algorithm for estimating pessimistic profit value ρ^- was developed:

```
Γ* = Γ
ξ = 0
ρ* = 0
while Γ* > 0 do
    ρ*_0 = ρ*
    for each k ∈ K do
        ξ' = ξ
        ξ'_k = ξ_k + min{ψ_k − ξ_k, Γ*}            ▷ Try cutting to the furthest possible extent
        X, y = solve Equation 2 with fixed perturbation ξ = ξ'
        ρ = Σ_{n∈N,d∈D} x_dn p_dn − c^T y
        if ρ < ρ* then
            ρ* = ρ
            k* = k
            ξ* = (Σ_{n∈N} a_kn y_n) − b_k           ▷ Reduce cut to the amount sufficient to obtain ρ
        end
    end
    Γ* = Γ* − (ξ* − ξ_k*)
    ξ_k* = ξ*
    if ρ* − ρ*_0 < ϵ then
        break                                       ▷ No significant improvement found
    end
end
return ρ*
```

Source: author's work.

The idea behind the heuristic presented in the above algorithm is to sequentially select the constraint that leads to the highest reduction of the cost function when under perturbation. At a given step of the algorithm, we are sequentially considering each constraint $k \in K$. For a given constraint k, we assign its maximum perturbation level so that it still satisfies the boxed-polyhedral uncertainty inequality presented in Equation (1) and solves the optimisation problem denoted by Equation (2). Finally, we choose a constraint whose perturbation leads to the highest reduction of the goal function presented in Equation (2). Once the new value is calculated, we set the perturbation level to eliminate the unnecessary slack. The algorithm stops when no significant improvement is found (the minimal improvement value is presented as $\epsilon > 0$ in the algorithm).

Figure 1. Dependencies between the proposed optimisation models



Source: author's work.

3. Numerical experiments

The goal of this section is twofold. Firstly, we will show the numerical accuracy of the algorithm from Section 2.2. Secondly, the sensitivity of the model's parameters will be demonstrated on sample input data.

The optimisation model presented in the previous section has been implemented in the Julia programming language. We conducted numerical experiments using Julia JuMP (Dunning et al., 2017; Legat et al., 2022).

The parameters for the numerical experiments are presented in Table 2. Please note that $\sim \{\dots\}$ indicates that a value is uniformly drawn from the given set, $\sim N(\mu, \sigma)$ means that a value is chosen from the normal distribution with mean μ and standard deviation σ, and $\sim U(a, b)$ denotes a number drawn from the uniform distribution. Random values are also used in the model in its equations – in this case they are shown in parentheses, e.g. $\max\left(0, (\sim N(4,7))\right)$ means a censored normal distribution where the negative values of the left tail are replaced with zeros. Finally, please note that whenever a continuous distribution is used, all the values are rounded to the nearest integer.

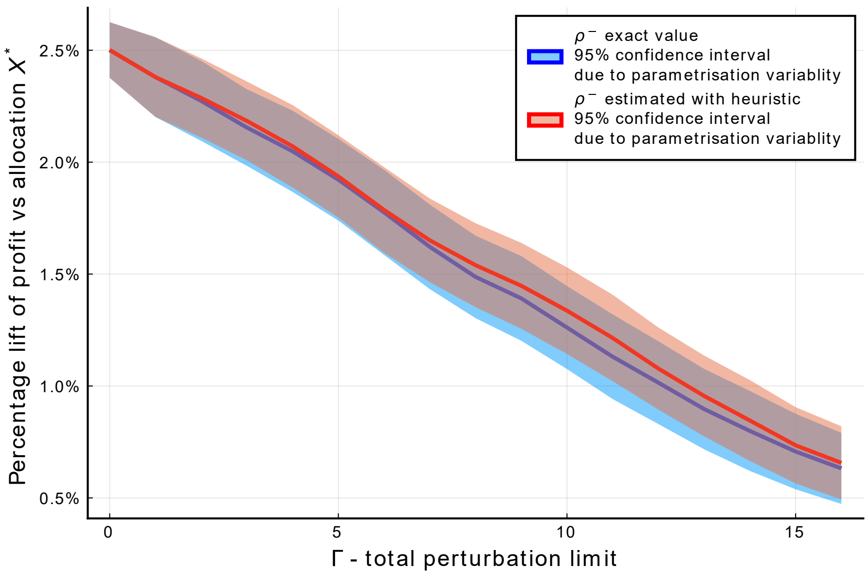
Table 2. Parameters used for numerical experiments

Parameter	Symbol	Value
Number of item types	$\|N\|$	15
Number of distributors	$\|D\|$	15
Number of critical part types	$\|K\|$	6
Substitution matrix	S	$s_{ii} = 1$ and $\forall_{i \neq j} P(s_{ij} = 1) = 1/3$
Technology matrix	a_{kn}	$\sim \{0,1\}$
Prices	p_{dn}	prices are generated for each value of $n \sim \{101 + (n-1) * \|D\|, \ldots, 100 + n * \|D\|\}$
Cost	c_n	$c_n = p_{1,n}/2$
Demand	v_{dn}	$\max\left(0, (\sim N(4,7))\right)$
Part availability	b_k	$\sum_{n \in N} \left(a_{kn} \sum_{d \in D} v_{dn} - (\sim U(1,20))\right)$
Maximum perturbation	ψ_k	$\sim \{4, 5, 6\}$
Total perturbation limit	Γ	1–16 (numerical accuracy) 5–40 (model properties)

Source: author's work.

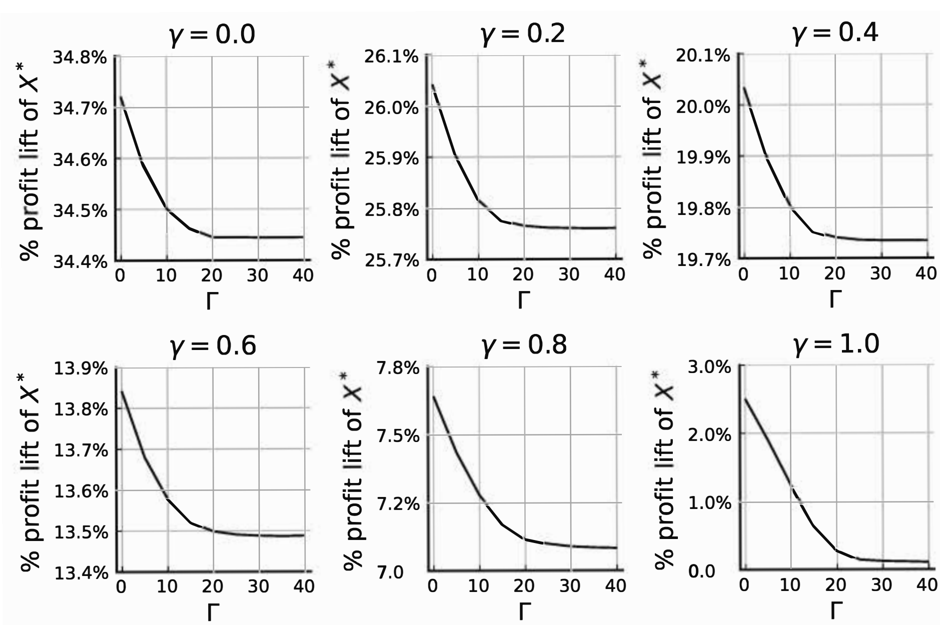
We start by evaluating the quality of the cutting plane heuristic proposed in the algorithm from Section 2.2. Following the parametrisations presented in Table 2, 30 different randomised scenarios were constructed, including demand structure, prices, substitution matrices and critical part availability. For each of those scenarios, the pessimistic value of profit ρ^- was evaluated in two ways. Firstly, ρ^- was fully enumerating all values $\boldsymbol{\xi} \in \mathbf{U}$ and solving a separate optimisation model, hence obtaining the exact solution to the robust optimisation problem presented in Equation (9). Secondly, the same ρ^- value was estimated with the algorithm. The results are presented in Figure 2, and scaled against the profit that can be obtained in the pessimistic scenario without substitution (X^*). The profit lift is calculated as $(\rho(X)/\rho(X^*) - 1) * 100\%$. It becomes evident that the heuristic yields a slightly larger estimate of profit ρ^- compared to the actual exact solution; however, this difference is marginal considering the influence of the other model parameters – any change of Γ has a much more significant impact on the results.

Figure 2. Performance of the algorithm developed in the paper vs. the exact solution



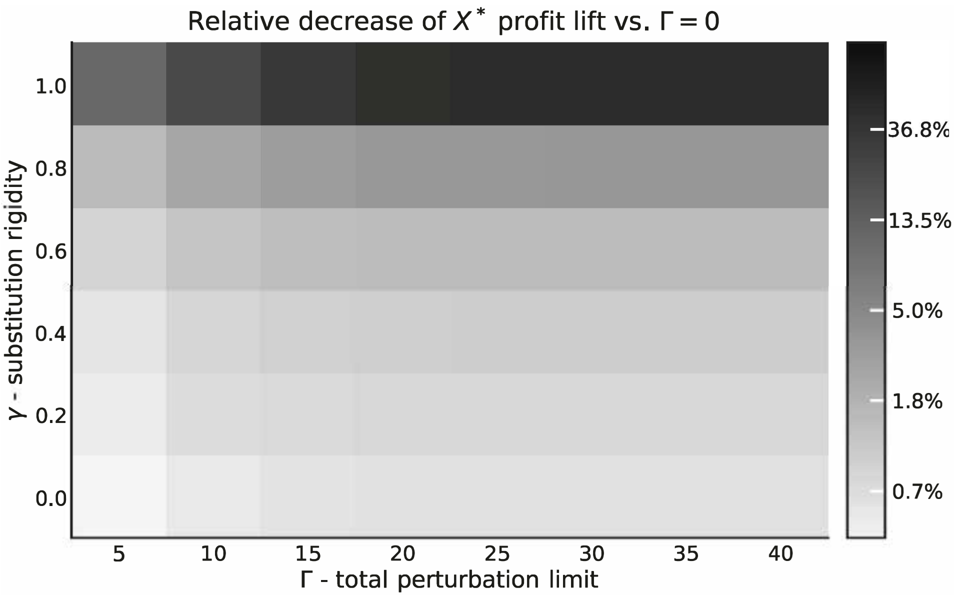
Source: author's calculations.

Figure 3. Additional pessimistic estimate of profits ρ^- acquired due to substitution at various total perturbation limits Γ and substitution rigidity levels γ



Source: author's calculations.

Figure 4. Impact on the overall profit of substitution rigidity and the total perturbation limit



Source: author's calculations.

Figure 3 shows how substitution rigidity γ jointly with total perturbation limit Γ influence the model outcomes. The profit that can be obtained at rigid pessimistic solution X^* is used as a benchmark. Similarly to the previous plot, the profit lift value is calculated as $(\rho(X)/\rho(X^*) - 1) * 100\%$. Figure 3 demonstrates that regardless of the substitution rigidity level γ, the pessimistic estimate of profits ρ^- drops with the increase of the total perturbation limit Γ, which is the expected outcome. However, it is worth noting that the marginal drops decrease as the values of Γ increase.

Figure 4 shows to what extent the presented model is sensitive to maximum perturbation level Γ at various levels of product substitution rigidity γ. Again, the profit lift $(\rho(X)/\rho(X^*) - 1)$ is used as the benchmark value (note the log scale of the colour bar). However, in order to ensure comparability, the results for each rigidity level γ have been scaled using the optimistic profit lift value, i.e. the profit lift without perturbation $(\Gamma = 0)$. It can be seen that when there is a high product substitution rigidity $(\gamma = 0)$, the model is very sensitive to the perturbation limit. On the other hand, when even small level of substitution is possible, the business importance of perturbation limit Γ quickly diminishes. This means that if customers are even slightly inclined to buy product substitutes on a market with shortages, then it immediately has a considerable impact on the number of goods that can be manufactured. Hence, even a small substitution flexibility yields a significant increase in profits.

The numerical experiments show that the heuristic proposed in the algorithm ensures a sufficient level of accuracy to apply the model in supporting a real production system. The model proposed in the paper applied using real-life data allows a more precise measurement of the effects that product substitutes have on the actual operational efficiency of the company. It also shows that ensuring even small substitution elasticity can have a significant influence on the financial results of a manufacturing enterprise.

4. Conclusions

In this paper, a robust optimisation model was presented which maximises the profits of a manufacturing company located in the middle of a supply chain. This kind of a company struggles with uncertain supplies of sub-components and experiences market disturbances, and therefore is willing to offer its customers substitutes instead of the originally requested products. The model developed in the paper was implemented in the Julia programming language and tested in a series of numerical experiments.

The main outcomes are as follows: (1) an integrated model for profit maximisation in a production company facing uncertain supplies of critical components at various levels of product substitution rigidity, (2) a heuristic that makes it possible to efficiently solve the presented problem at scale, (3) a set of business guidelines on how the product substitution rigidity and component availability perturbations affect the final financial situation of a company, and (4) managerial insights for decision-making in post-pandemic markets. The proposed results and methodology can be immediately applied to companies operating on today's markets prone to unbalanced demand and sub-component shortages.

The research presented in the paper can be expanded on in many ways. One of them is multi-period planning for resources, i.e. adding another dimension of weeks or months to the production plan. This would significantly increase the computational conditionality of the model. Another possible direction of the future research could involve the construction of an agent-based model (e.g. see Tesfatsion, 2003) of an entire shortage-driven economy. Such a model would take into consideration several manufacturers in the logistic chain, so that the output of one manufacturer would be the input for another.

Acknowledgements

This research was funded in whole by the National Science Centre, Poland (grant number: *2021/41/B/HS4/03349*). For the purpose of the Open Access policy, the author applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

References

Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust Optimization* (vol. 28). Princeton University Press. https://doi.org/10.1515/9781400831050.

Bertsimas, D., Dunning, I., & Lubin, M. (2016). Reformulation versus cutting-planes for robust optimization. *Computational Management Science*, *13*(2), 195–217. https://doi.org/10.1007/s10287-015-0236-z.

Beyer, H.-G., & Sendhoff, B. (2007). Robust optimization – A comprehensive survey. *Computer methods in applied mechanics and engineering*, *196*(33–34), 3190–3218. https://doi.org/10.1016/j.cma.2007.03.003.

Boston, W. (2022. March 3). Ukraine War Plunges Auto Makers Into New Supply-Chain Crisis. *Wall Street Journal*, *3*. https://www.wsj.com/articles/ukraine-war-plunges-auto-makers-into-new-supply-chain-crisis-11646309152.

Brakman, S., Garretsen, H., & van Witteloostuijn, A. (2020). The turn from just-in-time to just-in--case globalization in and after times of COVID-19. An essay on the risk re-appraisal of borders and buffers. *Social Sciences & Humanities Open*, *2*(1), 1–6. https://doi.org/10.1016/j.ssaho.2020.100034.

Cai, M., & Luo, J. (2020). Influence of COVID-19 on Manufacturing Industry and Corresponding Countermeasures from Supply Chain Perspective. *Journal of Shanghai Jiaotong University (Science)*, *25*(4), 409–416. https://doi.org/10.1007/s12204-020-2206-z.

Chen, J., Wang, H., & Zhong, R. Y. (2021). A supply chain disruption recovery strategy considering product change under COVID-19. *Journal of Manufacturing Systems*, *60*, 920–927. https://doi.org/10.1016/j.jmsy.2021.04.004.

Dunning, I., Huchette, J., & Lubin, M. (2017). JuMP: A Modeling Language for Mathematical Optimization. *SIAM review*, *59*(2), 295–320. https://doi.org/10.1137/15M1020575.

Legat, B., Dowson, O., Garcia, J. D., & Lubin, M. (2022). MathOptInterface: A Data Structure for Mathematical Optimization Problems. *INFORMS Journal on Computing*, *34*(2), 672–689. https://doi.org/10.1287/ijoc.2021.1067.

Li, X., Wang, B., Liu, C., Freiheit, T., & Epureanu, B. I. (2020). Intelligent Manufacturing Systems in COVID-19 Pandemic and Beyond: Framework and Impact Assessment. *Chinese Journal of Mechanical Engineering*, *33*(1), 1–5. https://doi.org/10.1186/s10033-020-00476-w.

Li, Z., Ding, R., & Floudas, C. A. (2011). A comparative theoretical and computational study on robust counterpart optimization: I. robust linear optimization and robust mixed integer linear optimization. *Industrial & Engineering Chemistry Research*, *50*(18), 10567–10603. https://doi.org/10.1021/ie200150p.

Mbah, R. E., & Wasum, D. F. (2022). Russian-Ukraine 2022 war: A Review of the Economic Impact of Russian-Ukraine Crisis on the USA, UK, Canada, and Europe. *Advances in Social Sciences Research Journal*, *9*(3), 144–153. https://doi.org/10.14738/assrj.93.12005.

Paul, S. K., & Chowdhury, P. (2021). A production recovery plan in manufacturing supply chains for a high-demand item during COVID-19. *International Journal of Physical Distribution & Logistics Management*, *51*(2), 104–125. https://doi.org/10.1108/IJPDLM-04-2020-0127.

Queiroz, M. M., Ivanov, D., Dolgui, A., & Fosso Wamba, S. (2020). Impacts of epidemic outbreaks on supply chains: mapping a research agenda amid the COVID-19 pandemic through a structured literature review. *Annals of Operations Research*, 1–38. https://doi.org/10.1007/s10479-020-03685-7.

Shahed, K. S., Azeem, A., Ali, S. M., & Moktadir, M. A. (2021). A supply chain disruption risk mitigation model to manage COVID-19 pandemic risk. *Environmental Science and Pollution Research*, 1–16. https://doi.org/10.1007/s11356-020-12289-4.

Singh, S., Kumar, R., Panchal, R., & Tiwari, M. K. (2021). Impact of COVID-19 on logistics systems and disruptions in food supply chain. *International Journal of Production Research*, *59*(7), 1993–2008. https://doi.org/10.1080/00207543.2020.1792000.

Tesfatsion, L. (2003). Agent-based computational economics: modeling economies as complex adaptive systems. *Information Sciences*, *149*(4), 262–269. https://doi.org/10.1016/S0020-0255(02)00280-3.

Tsolas, S. D., & Hasan, M. M. F. (2021). Survivability-aware design and optimization of distributed supply chain networks in the post COVID-19 era. *Journal of Advanced Manufacturing and Processing*, *3*(3), 1–14. https://doi.org/10.1002/amp2.10098.

Przegląd Statystyczny. Statistical Review, 2022, vol. 69, 2, 37–39
Received: 27.08.2022, revised: 1.09.2022, accepted: 10.09.2022

In memory of Professor Stanisława Bartosiewicz

Krzysztof Jajuga[a]

Stanisława Maria Helena Bartosiewicz was born on 8 May 1920 in Brzeżany (Tarnopolskie Voivodship). In 1938, she graduated from a neoclassical gymnasium and began studying at the Academy of Foreign Trade in Lviv, which was then part of Poland. After the outbreak of the war, Stanisława Bartosiewicz suspended her studies and began to work. In 1946, she settled in Lower Silesia. While still working, in 1947 she resumed her studies at the newly established private Higher School of Commerce in Wrocław. In 1949, Stanisława Bartosiewicz received a graduation diploma of the 1st degree. In 1953, she obtained a master's degree in economics in the field of statistics. Her master's thesis was entitled *Regression analysis as a tool for assessing the profitability of enterprises* (Pol. Analiza regresji jako narzędzie badania gospodarności przedsiębiorstw). In 1962, she earned a PhD in economic sciences. Her doctoral thesis was entitled *Adequacy of indicators characterising the activity of enterprises* (Pol. Adekwatność wskaźników charakteryzujących działalność przedsiębiorstw). She obtained a habilitation degree (i.e. postdoctoral degree) in economics in 1984 and in 1988 she became professor of economic sciences.

Professor Bartosiewicz co-founded the Wrocław School of Econometrics. The most important areas of her research included econometric modelling, multi-

[a] Wroclaw University of Economics and Business, Faculty of Economics and Finance, ul. Komandorska 118/120, 53-345 Wrocław, Poland, e-mail: krzysztof.jajuga@ue.wroc.pl,
ORCID: https://orcid.org/0000-0002-5624-6929.

dimensional comparative analysis and mathematical methods of decision-making. She devised numerous methods in the above areas, the most acknowledged of which were the graph method of selecting explanatory variables in linear and nonlinear models and the method of selecting the analytical form of a model with many explanatory variables.

Professor Bartosiewicz wrote two monographs relating to the field of econometric modelling. *Econometrics. Technology of econometric information processing* (Pol. Ekonometria. Technologia ekonometrycznego przetwarzania informacji) presents the full sequence of processing economic information using an econometric model. *Specification of econometric models and their use in the analysis of socio-economic phenomena* (Pol. Specyfikacja modeli ekonometrycznych i ich wykorzystanie w analizie zjawisk społeczno-gospodarczych) is a summary and review of Professor Bartosiewicz's original scientific achievements. The publication covers all the issues concerning the process of econometric modelling and inference.

Combining the theory of econometric modelling with economic applications, especially on a micro-scale was a very important aspect of the scientific activity of Professor Bartosiewicz. In her scientific work, she always strove for the transparency, communicativeness and practicality of her theoretical considerations. She had the rare ability to recognise the potential of a practical approach and quickly interpret complex quantitative methods. She had excellent analytical skills complemented by an exceptional ability to generalise accurately.

Professor Bartosiewicz's achievements in educational activity were equally striking. Several distinguished researchers were once her students. She managed to equip them with knowledge and motivated them to continue and develop her academic legacy. Her unique approach to teaching favoured focusing on an individual and involved developing her students' scientific personalities. She encouraged others to unleash their creativity and scientific intuition. Professor Bartosiewicz's credo was: don't talk about the technical aspects of things, focus on the idea. She also had a rare ability to conduct conversations in a way that led to solving problems.

She brought together the community of Polish econometricians and statisticians, from doctoral students to professors, who could draw from her vast experience and knowledge during numerous conferences, e.g. the conference of the Econometricians of Southern Poland, conferences in Zakopane, and many others. For several terms, Professor Bartosiewicz was a member of the Statistics and Econometrics Committee of the Polish Academy of Sciences.

Professor Bartosiewicz never stopped striving for self-development. She was modest about her accomplishments and spoke about them lightly, e.g. when at the age of 84 she started to learn how to use a computer. She used to write structured poems based on Japanese patterns. In her private life, she was known for her great sense of humour and being an expert on borderland jokes, told with great eloquence and in original accent. She also wrote a humorous book on complex issues related to econometrics, entitled *Econometrics with a pinch of salt* (Pol. Ekonometria z przymrużeniem oka).

Professor Bartosiewicz once shared a set of simple, yet valuable and universal pieces of advice:

How to be happy? A few inspirational thoughts about human happiness:

1. If you want to be happy, then learn;
2. If you want to be happy, be tolerant of your environment;
3. If you want to be happy, learn to assess your own qualities in relation to those of others;
4. If you want to be happy, have a good sense of humor;
5. If you want to be happy, always bear in mind that constant success, great achievements, praise and rewards from others lead to boredom.

Professor Bartosiewicz passed away in Wrocław on July 21, 2022.
She will forever remain in our memories.