

DNB Working Paper

No. 624 / January 2019

Outlier detection in TARGET2 risk indicators

Ronald Heijmans and Chen Zhou

DeNederlandscheBank

EUROSYSTEEM

Outlier detection in TARGET2 risk indicators

Ronald Heijmans and Chen Zhou *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 624

January 2019

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

Outlier detection in TARGET2 risk indicators *

Ronald Heijmans^{a,b} and Chen Zhou^{a,c}

^a*De Nederlandsche Bank*

^b*Tilburg University*

^c*Erasmus University*

January 2019

Abstract

This paper studies the detection of outliers in risk indicators based on large value payment system transaction data. The ten risk indicators are daily time series measuring various risks in the large value payment system, such as operational risk, concentration risk and liquidity flows related to other financial market infrastructures. We use extreme value theory and local outlier factor methods to identify anomalous data points (outliers). In a univariate setup, the extreme value analysis quantifies the unusualness of each outlier. In a multivariate setup, the local outlier factor method identifies outliers by measuring the local deviation of a given data point with respect to its neighbours. We find that most detected outliers are at the beginning and near end of the calendar month when turnover is significantly larger than at other days. Our method can be used e.g. by overseers and financial stability experts who wish to look at many (risk) indicators in relation to each other.

Keywords: risk indicator, TARGET2, financial market infrastructure, extrem value theory (EVT), local outlier factor (LOF), anomaly.

JEL classifications: E42, E50, E58, E59.

*Heijmans and Zhou can be reached at ronald.heijmans@dnb.nl and chen.zhou@dnb.nl, respectively. Heijmans and Zhou are members of one of the user groups with access to TARGET2 data in accordance with Article 1(2) of Decision ECB/2010/9 of 29 July 2010 on access to and use of certain TARGET2 data. DNB and the MIPC have checked the paper against the rules for guaranteeing the confidentiality of transaction-level data imposed by the MIPC pursuant to Article 1(4) of the above mentioned issue. The views expressed in the paper are solely those of the authors and do not necessarily represent the views of the Eurosystem or De Nederlandsche Bank.

1 Introduction

Financial market infrastructures (FMIs) play a crucial role in the well functioning of the economy. They facilitate the clearing and settlement of financial obligations. Most of these obligations are the consequence of economic activity (buying and selling goods and services). Besides, economic activity, many (large) payment flows are the consequence of monetary policy and regulatory requirements (such as reserve maintenance and balance sheet requirements). When financial actors (such as banks, insurance companies, national treasury departments etc.) are not able to fulfill their obligations in time, it can seriously damage the economy. Therefore, FMIs have to meet high international standards laid down in the Principles for Financial Market Infrastructures (PFMIs, CPSS (2012)).¹

Berndsen and Heijmans (2017) developed risk indicators for the most important large value payment system of the European Union for euro denominated payments (TARGET2). They have converted transaction data into risk indicators providing quantitative information on the risks defined by these PFMIs. To determine whether a risk indicator shows unusual behavior (i.e. can be seen as an outlier) they use the univariate method of Timmermans et al. (2018).² This method is an ARIMA model which corrects for weekly and monthly cyclical patterns. The model compares the forecasted value with the actual observations. When the actual value lies outside e.g. 99% interval of the forecasted value it is considered an outlier.³

This paper extends the outlier detection study in Berndsen and Heijmans (2017) by applying two statistical tools for identifying univariate and multivariate outliers in the risk indicators, namely, extreme value theory and the local outlier factor method. In the univariate setup, extreme value theory helps to determine how unlikely a new observation is, relative to its history. The unusualness is quantified as once in how many years such an outlier would occur. In case the number of years is large (e.g. more than one year, or five years) it is considered an outlier. The actual threshold (one or five years) can be set afterwards by the expert in charge. Compared to Berndsen and Heijmans (2017), the extreme value theory approach does not assume that the forecasting error follows a normal distribution; instead the historical data determine the tail of the distribution.

¹CPSS is nowadays called CPMI, i.e. Committee on Payments and Market Infrastructures. The Principles are international standards for 1) payment systems that are systemically important, 2) central securities depositories, 3) securities settlement systems, 4) central counterparties and 5) trade repositories.

²Their work is closely related to the detection of cyclical patterns in large value payment system data by Van Ark and Heijmans (2016).

³The difficulty is that the values of the indicator are not normally distributed. Therefore, they choose to look at multiple times the real value is outside the 95% and 99% interval.

In addition to the univariate analysis, we also detect multivariate outliers by combining several risk indicators. Univariate outliers are always extreme values in the time series, either unusually high or unusually low. By contrast, multivariate outliers can be an unusual combination of values of risk measures on a given day. A multivariate outlier may actually correspond to moderate levels at all marginals. The unusualness is solely attributed to the combination. We employ the local outlier factor method for multivariate outlier detection. The local outlier factor method identifies outliers by measuring the local deviation of a given data point with respect to its neighbours. We show that the two new outlier detection methods can be used by e.g. overseers and operators to get a better understanding of the large payment system and of changes therein.

Our paper adds to the growing literature in defining (risk) indicators and detecting outliers in FMIs. A large part of the large value payments system literature looks at the timing of payments. Especially free riding behavior of participants is extensively studied. Free riding is the intentional delay of payments in which the participant waits to make its own payments until it has received liquidity. Bech and Garratt (2003) developed a theoretical framework of payments delay.⁴ Diehl (2013) studied the presence of free riding of German banks in TARGET2. Massarenti et al. (2012) developed an indicator for timing of the whole TARGET2 system. Baek et al. (2014) developed an indicator to measure the state of liquidity in the Korean large value payment system, BoK-wire. This measure provides information on whether the system has sufficient liquidity to fulfil all forecasted payment obligations by the end of the day. Heijmans and Heuver (2014) investigated payments' delay and identified key elements of liquidity stress for individual banks in TARGET2. They developed a method to monitor the banks in TARGET2. Kaliontzoglou and Müller (2015) zoomed in to aspects related to payments timing indicators. Li and Pérez-Saiz (2016) measured systemic risk in the network of FMIs as the probability that two or more FMIs have a large credit risk exposure to the same FMI participants. They construct indicators for three main Canadian FMIs using extreme value methods to estimate the probability. Squartini et al. (2013) showed early-warning signals for topological collapse in interbank networks. Heijmans et al. (2016) created a dynamic way of looking at changes in the topological structure of payment's data. They create a movie showing the changes in the network structure dynamically. A different method to detect outliers in payments' data is by Triepels et al. (2018). They used a neural network (auto-encoder) to detect anomalous payment flows between commercial banks

⁴Abbink et al. (2017) and Heemeijer and Heijmans (2015) conducted a lab experiment based on their game theoretical model.

in a large value payment system. Their method works well to detect bank runs, which show more than usual outgoing payment flows.

This paper also contributes to the literature on outlier detection, both for univariate and multivariate outlier detection. For univariate outlier detection, we apply extreme value theory. Although extreme value theory has been developed for more than 60 years, see e.g. textbooks such as Embrechts et al. (1997) and de Haan and Ferreira (2006), its application to outlier detection has a shorter history of about 20 years since the work of Roberts (1999). This method has been applied to, for example, biomedical data (Roberts (2000)), damage detection and control (Sohn et al. (2005)) and aircraft engine health monitoring (Sundaram et al. (2009)).

For multivariate outlier detection, we apply the local outlier factor method, which was initially proposed in Breunig et al. (2000). Since then, this method has been applied to various fields, for example, for building the Intrusion Detection System (IDS) for computer network security (Loureiro et al. (2004)). Overall, our paper brings the two outlier detection methods to a new field: payment system monitoring. It shows that they help to identify irregularities in a large payment system.

The organisation of the paper is as follows. Section 2 describes the TARGET2 system and data. Section 3 describes the construction and statistics of the individual risk indicators and the univariate extreme value theory method. Section 4 explains the multi-variate analysis. Section 5 provides a theoretical framework. Section 6 summarizes and provides policy conclusions.

2 TARGET2

2.1 The system TARGET2

TARGET2 is the real-time gross settlement (RTGS) system owned and operated by the Eurosystem.⁵ Payment transactions in TARGET2 are settled one by one on a continuous basis, in central bank money with immediate finality. There is no upper or lower limit on the value of payments.

TARGET2 settles payments related to monetary policy operations, interbank and customer payments, and payments related other FMIs handling the euro (such as Euro1, CLS, securities settlement systems or central counterparties). Most of the 1,076 direct participants are euro area credit institutions,

⁵TARGET2 stands for Trans-European Automated Real-time Gross settlement Express Transfer system. TARGET2 has been introduced in November 2007. The countries participating in TARGET2 migrated to TARGET2 in three waves, from November 2007 to May 2008.

several large non-euro area banks (notably UK, Swiss and US) and euro area central banks.⁶

Table 1 provides general statistics on TARGET2 for the year 2016. It processes on average more than 340,000 transactions daily with a corresponding turnover of 1,700 billion euros. The average transactions value is EUR 5 million.

Table 1: TARGET2 facts in 2016.

Number of participants	TARGET2 had 1,076 direct participants, 701 indirect participants and 5,072 correspondents
Number of ancillary systems	TARGET2 settled the cash positions of 80 ancillary systems
Daily averages	TARGET2 processed a daily average of 342,008 payments, representing a daily average value of EUR 1.7 trillion
Average transaction value	EUR 5 million
Payment values	More than 70% of all TARGET2 payments had a value of less than EUR 50,000 each: 10% of all payments had value of over EUR 1 million
Large-value payment system traffic	TARGET2's share in total large-value payment system traffic in euro was 90% in value terms and 63% in volume terms

source: ECB (<https://www.ecb.europa.eu/paym/target/target2/facts/html/index.en.html>)

2.2 TARGET2 data

Our data set contains all individual transactions settled in TARGET2 from January 2010 until December 2016. TARGET2 distinguishes between several payment types. The payments are divided into four different groups: 1) main interbank transactions, 2) payments with a central bank involved, 3) payments from and to ancillary systems (AS) and 4) liquidity transfers. Ancillary systems are other FMIs settling in TARGET2. Liquidity transfers between the accounts of the same bank (category 4.4), is not seen as an actual payment, but just an administrative shift of liquidity between two accounts of the same legal entity. Therefore, category 4.4 is often omitted from the published ECB TARGET2 statistics as these transactions do not reflect an economic process. For this reason, we also exclude category 4.4 from our analysis. Categories 4.6 and 4.7 have been introduced at the start of TARGET2 securities (T2S) in 2015.⁷ Although in this paper we focus on all TARGET2 transaction (excluding

⁶Besides the euro area countries also Bulgaria, Croatia, Denmark, Poland and Romania have access to TARGET2 (status December 2018) have access to TARGET2. At the start of our data set there have been a few countries in the European Union that did not have the euro yet: Estonia, Latvia, Lithuania, Slovakia.

⁷The migration of T2S is also done in migration waves. This payment category will therefore increase over time.

category 4.4) it is also possible to look at a single or a group of payment types.⁸

3 Data description and univariate analysis

In this section, we first describe how the risk indicators time series have been constructed (section 3.1). Next, in Section 3.2 we explain the procedures to select the optimal time series model for filtering each series. Based on the filtered errors, we detect outliers for each individual indicator time series and evaluate the unusualness of detected outliers using extreme value theory in Section 3.3.

3.1 Construction of the risk indicator times series

The risk indicators we use in are based on the ones developed by Berndsen and Heijmans (2017). These indicators identify quantitative risks in TARGET2, which are inspired by the Principles for Financial Market Infrastructures defined by CPSS IOSCO (see CPSS, 2012). Berndsen and Heijmans (2017) focus on operational risk, concentration risk and liquidity (inter) dependencies. These indicators provide information on the system at daily frequency.⁹

Table 2 gives an overview of the ten indicators we have taken from Berndsen and Heijmans (2017). Appendix A shows the time series of each of the indicators. In the second column the description of the indicator is given and in the third column the type of Principle the indicator provides quantitative information on. We use this set of indicators to set up a more advanced outlier detection method than Timmermans et al. (2018).

The relative use of the operational capacity is the amount of transactions processed in TARGET2 relative to the maximum the system guarantees. The throughput guideline shows the percentage of total payments value settled before a certain time (here 12.00 hours). The Herfindahl Hirschman index (HHI, indicator 3 to 8) is a concentration measure of the system. The concentration is measured at the participant level and looks at concentration of the system. The HHI outgoing and total turnover take the concentration of the outgoing and total payment flows (indicator 3 and 4), respectively. Indicators 5 to 8 look at the concentration in terms of network characteristics (degree, eigenvector centrality, hub centrality and authority centrality). Indicator 9 presents the relative share of payments made to other FMIs in TARGET2 relative to the total TARGET2 turnover. The last indicator (10) shows the

⁸See Berndsen and Heijmans (2017) for more information on the payment types.

⁹It is also possible to look at other frequencies, such as hourly or monthly.

Table 2: List of indicators.

Indicator Number	Indicator description	PFMI risk
1	Relative use of operational capacity	Operational
2	Throughput guideline	Operational
3	HHI outgoing turnover	Concentration
4	HHI total turnover	Concentration
5	HHI degree (unweighted)	Concentration
6	HHI eigenvector centrality (unweighted)	Concentration
7	HHI hub centrality	Concentration
8	HHI authority centrality	Concentration
9	Turnover to other FMIs	Interdependencies
10	Net bilateral flows	Interdependencies

net bilateral flows, which means the remaining liquidity flows between banks if all bilateral flows between participants are netted.

Table 3 provides general statistics on the ten indicators mentioned in Section 3.1. The table shows the 10th to 90th percentile and the auto correlation with lag 1 of these ten indicator time series. From these percentiles, we observe that the indicators are not normally distributed, which is in line with the findings in Timmermans et al. (2018). In addition, Indicators 6 and 7 show a high degree of auto correlation, while indicators 2 to 4 and 8 to 10 have relatively weak correlation. Indicators 1 and 5 have medium degree of correlation. Since most indicators have some to strong auto correlation, we shall employ univariate time series models to filter out the serial dependence.

3.2 Time series filtering

For each indicator of Table 2, we use a time series model to filter out its serial dependence. After the filtering process, we assume that the filtered series form an independent and identically distributed (i.i.d.) series. We use the full sample for the optimal model selection. However, the selected optimal model will be re-estimated in each estimation window. For each given indicator $\{Y_t\}$, since they are all percentages between zero and one, we first conduct a logit transformation as

$$\tilde{Y}_t = \log(Y_t/(1 - Y_t)),$$

Table 3: Risk indicator statistics based on data between 2010-2016.

Indicator number	10–percentile	first quartile	median	third quartile	90–percentile	auto correlation
1	0,46	0,5	0,55	0,60	0,67	0,59
2	0,41	0,44	0,47	0,50	0,53	0,46
3	0,0144	0,0152	0,016	0,017	0,0185	0,28
4	0,0139	0,0146	0,0152	0,016	0,0168	0,20
5	0,0034	0,0035	0,0036	0,0037	0,0037	0,56
6	0,0867	0,0899	0,1041	0,1137	0,1224	0,94
7	0,0024	0,0025	0,0025	0,0026	0,0028	0,88
8	0,0029	0,003	0,0031	0,0032	0,0033	0,35
9	0,2148	0,2339	0,2576	0,3007	0,3615	0,37
10	0,39	0,41	0,43	0,45	0,47	0,46

such that $\{\tilde{Y}_t\}$ is on the entire real line. For indicator 9, since the series shows a trend, we consider its day-to-day difference. For indicator 3, since only a low value is of concern for operational risk, we multiply the series by -1 . After these transformations, we obtain a time series $\{X_t\}$ for each indicator. Without loss of generality, we simply use the notation $\{X_t\}$ to indicate any (transformed) indicator.

We fit an ARIMA(p, d, q) model with covariates to $\{X_t\}$ as follows

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (1 - L)^d X_t = \delta + \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_t + \beta' Z_t,$$

, where L stands for the lag operator, $\{\varepsilon_t\}$ is an i.i.d. innovation series, $\{Z_t\}$ is the set of covariates at time t . In this analysis, we only consider three dummy variables indicating whether day t is the end of a month, a quarter or a year as potential covariates, denoted as $IM_t = 1_{\{\text{day } t \text{ is the end of a month}\}}$, $IQ_t = 1_{\{\text{day } t \text{ is the end of a quarter}\}}$ and $IY_t = 1_{\{\text{day } t \text{ is the end of a year}\}}$.

With three potential covariates, we consider eight models corresponding to all different subsets of these covariates: \emptyset , $\{IM_t\}$, $\{IQ_t\}$, $\{IY_t\}$, $\{IM_t, IQ_t\}$, $\{IQ_t, IY_t\}$, $\{IM_t, IY_t\}$ and $\{IM_t, IQ_t, IY_t\}$. For each given model, we select the optimal (p, d, q) by minimizing the Akaike Information Criterion (AIC). Then across the eight best models, we select the one that corresponds to the lowest AIC value. The selected model consists of the optimal p^*, d^*, q^* and the optimal subset of the dummies Z^* . We remark that this process is only run once and only the best model structure is used.

Next, for each date t in the analysis period, we fit the optimal model to the $n = 1250$ days proceeding date t . In other words, we fit an $\text{ARIMA}(p^*, d^*, q^*)$ model with covariates $\{Z_t^*\}$ to $X_{t-n}, X_{t-n+1}, \dots, X_{t-1}$ and obtain the estimated coefficients $\hat{\phi} = (\hat{\phi}_1, \dots, \hat{\phi}_{p^*})'$, $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_{q^*})'$, $\hat{\delta}$ and $\hat{\beta}$. In addition, we obtain the filtered innovation terms $\{\hat{\epsilon}_s\}_{s=t-n}^{t-1}$.

We use the estimated coefficients, innovations and Z_t^* to forecast the indicator at time t as \hat{X}_t . Then we define $\hat{\epsilon}_t = X_t - \hat{X}_t$ as the forecasting error. Assuming that the indicator series follows the ARIMA model with the chosen covariate, then $\{\epsilon_s\}_{s=t-n}^t$ is an i.i.d. sample of $n+1$ random errors from a common distribution F . Therefore, we can use the n estimated innovations $\{\hat{\epsilon}_s\}_{s=t-n}^{t-1}$ to estimate the distribution F , and then consider where $\hat{\epsilon}_t$ is an outlier of the estimated distribution.

3.3 Univariate outlier detection: extreme value analysis

Since for most cases, we are interested in the right tail of the indicator series only, i.e. only an extremely high value of the indicator is of policy concern, we use the probability

$$p_t = 1 - F(\hat{\epsilon}_t)$$

as the outlier probability for the forecasting error. If the outlier probability is extremely low, we conclude that the forecasting error on day t is exceptionally high, or in other words, the deviation of the observation X_t from the optimal ARIMA model is extremely unlikely. In that case, we identify date t as an outlier day. Further, we convert the probability to a frequency as $1/p_t/250$. Here, we regard the number of operational days in a year as 250. The frequency calculated indicates once per how many years such an exceptional day may occur. If the frequency is above a certain threshold, say 1 or 5, one can define the day as an outlier. Note that since we calculate the frequency at each day, the standard to identify an outlier can be determined ex post by a policy maker.

If we are also interested in the left tail of the indicator series, i.e. an extremely low value of the indicator is also of policy concern, then we can use the probability $F(\hat{\epsilon}_t)$. This case can easily be handled by working with the negative series $\{-\hat{\epsilon}_s\}_{s=t-n}^t$. For simplicity, we will only discuss the right tail case.

With the time series filtering, the outlier detection issue boils down to a simple statistical question. Suppose we have i.i.d. observations $\{Y_s\}_{s=1}^n$ drawn from a common distribution function F . Suppose we have a new observed value y , what is the probability that y is an outlier.

If y is not exceptionally high, then the probability can be easily estimated by its empirical counter parts

$$\hat{p}_{emp} = \frac{1}{n} \sum_{i=1}^n 1_{Y_i > y} \quad (1)$$

However, since we intend to capture outliers with very low probability, we will handle an exceptionally high value of y . If $y > \max_{s=1}^n Y_s$, then $\hat{p}_{emp} = 0$ for the given sample $\{Y_s\}_{s=1}^n$. It is necessary to be able to estimate p in such a situation. Further, even for a not-so-high value of y such that \hat{p}_{emp} reaches a non-zero but low level such as $1/n$, $2/n$, it is clear that the empirical estimate is inaccurate in such a situation. Overall, we intend to handle the case where $y = y_n$ such that $n(1 - F(y_n)) \rightarrow \lambda$ with $\lambda \geq 0$ a fixed constant. In such a case, the empirical estimate fails and we opt to use Extreme Value Theory (EVT) in order to find more accurate estimator.

EVT assumes that the tail of the distribution F possesses some regularity that can be used for extrapolation. More specifically, if F has no finite endpoint, the so-called domain of attraction condition assumes that

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma},$$

where γ is the extreme value index; see, e.g. Embrechts et al. (1997).

We make an extrapolation in the tails following the domain of attraction condition as follows. Denote $Y_{1,n} \leq \dots \leq Y_{n,n}$ as the orderstatistics of $\{Y_s\}_{s=1}^n$. For an intermediate sequence $k := k(n)$ such that $k/n \rightarrow 0$, $k \rightarrow \infty$ as $n \rightarrow \infty$, the orderstatistic $Y_{n-k,n}$ can be viewed as an estimator of the $(1 - k/n)$ -quantile of F . Therefore

$$\frac{1 - F(y_n)}{k/n} \approx \frac{1 - F(y_n)}{1 - F(Y_{n-k,n})} \left(\frac{y_n}{Y_{n-k,n}} \right)^{-1/\gamma}.$$

Consequently, we obtain the EVT estimator of $p_n = 1 - F(y_n)$ as

$$\hat{p}_n := \frac{k}{n} \left(\frac{Y_{n-k,n}}{y_n} \right)^{1/\hat{\gamma}}, \quad (2)$$

where $\hat{\gamma}$ is a proper estimator of the extreme value index; see Weissman (1978). For high threshold y_n the EVT estimator is superior to the empirical estimator.

For the outlier detection analysis, we use the Hill estimator $\hat{\gamma}_H$ as the estimator of the extreme value

index (Hill (1975)), defined as follows:

$$\hat{\gamma}_H = \frac{1}{k} \sum_{i=1}^n \log Y_{n-i+1,n} - \log Y_{n-k,n}.$$

In our outlier detection analysis, the input observations $Y_s = \hat{\epsilon}_{t-s}$ for $s = 1, 2, \dots, n$. The tested threshold is $y = \hat{\epsilon}_t$. We first apply the empirical estimator \hat{p}_{emp} in (1) to estimate $p = 1 - F(y)$. If $\hat{p}_{emp} \geq k/n$ we keep it as the estimate, since the threshold y is not in the tail. If $\hat{p}_{emp} < k/n$, we regard the threshold y as a high threshold and consequently apply the EVT estimator in (2). One key step in the EVT estimator is to choose the tuning parameter k . In this analysis, we use an estimation window with a length of 1250 days, i.e. $n = 1250$. Correspondingly, we choose $k = 50$ such that $k/n = 4\%$.

Table 4 presents the outliers of the univariate analysis. The table shows which indicator had an outlier score of larger than 1 year and at which dates. From the table we can see that most of the outliers are found close to the end or the beginning of a new month (quarter or year). As we do correct for end of month (quarter and year) effect, the dates found here deviate substantially from their “normal” behaviour at these dates. Besides, there are some outliers after a public holiday (Pentecost) or on a public holiday that is not a TARGET2 closing day (Ascension day).

Table 4: List of univariate outliers.

Number	Date	Outlier score	Comment
1	March 29, 2016	1.32	close to end of months
3	May 25, 2015	3.33	day after public holiday (Pentecost)
	May 14, 2015	1.22	Ascension day (no TARGET2 closing)
4	May 25, 2015	3.66	day after public holiday (Pentecost)
	May 14, 2015	1.40	Ascension day (no TARGET2 closing)
5	August 15, 2016	1.07	
6	March 29, 2016	5.06	close to end of months
	April 7, 2015	3.87	
7	May 25, 2015	5.82	day after public holiday (Pentecost)
	August 31, 2015	2.15	end of month
	May 30, 2016	1.91	close to end of months
9	January 6, 2015	2.08	beginning of month
	June 29, 2016	1.41	close to end of 2nd quarter

note 1: The number refers to the indicator number defined in Table 2.

note 2: The outlier score gives how many times such a value is seen once in q years. Only scores above $q = 1$ are shown here.

4 Multivariate analysis

4.1 Local outlier factor

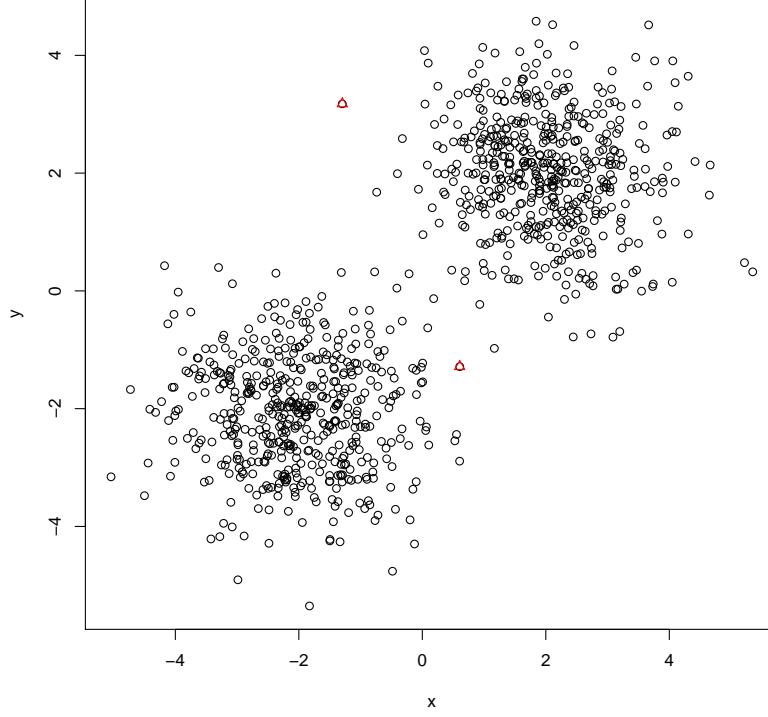
The individual series analysis outlined in Section 3.3 can quantify the unusualness of the outliers for each series. Notice that the outlier for an individual series always corresponds to an extreme value of the underlying indicator: either a very high or very low value of the observed indicator. The essence of outlier detection for a single series remains the same when considering the time series filtering approach: the time series model provides a forecast for the indicator in the next period, an outlier is therefore corresponding to a high or low “surprise” compared to the predicted value.

In practice, when considering multiple series, an outlier does not have to correspond to an extreme value of either series. For instance, consider an artificial dataset shown in Figure 1 based on two hypothetical series x and y . From the scatter plot, we observe two clusters for the data at the top-right and bottom-left corner. This leaves an observation in the middle of the scatter as an outlier, since it does not belong to any of the two clusters; for example the two data points indicated by a red triangular. Such an outlier is called the “local outlier”. To detect local outliers in multivariate series, we need a different method compared to the extreme value theory used for individual series.

The local outlier factor (LOF) method is a method specifically designed for detecting the local outliers, see Breunig et al. (2000). It assigns a score to each data point which measures the local deviation with respect to its neighbours. The data points with the highest scores are considered as local outliers. To illustrate the method in our context, we consider two indicator series HHI outgoing and total outgoing turnover (indicator 3 and 4) and net bilateral flows (NBF, indicator 10), denoted as $\{(X_t, Y_t)\}$, and perform an in-sample exercise using the LOF method.

Following the time series filtering as described in Section 3.2, we first identify the optimal $ARIMA(p^*, d^*, q^*)$ model with covariates for each marginal series using the full sample $t = 1, 2, \dots, T$. With the estimated model parameters, we filter out the corresponding innovation series for each series, denoted as $\{(\hat{\epsilon}_t, \hat{\xi}_t)\}_{t=1}^T$. Figure 2 shows the scatter plots of the filtered innovation series. From Figure 2a, we observe that there are two concentrated scenarios: a main scenario corresponding to high HHI total turnover and low net bilateral flows and a secondary scenario corresponding to low HHI total turnover and high net bilateral flows. If either case occurs, it would not be recognized as an outlier. The goal of the LOF method is to identify local outliers from these scatters.

Figure 1: An illustrative example for local outlier.



We first define k -distance and k -distance neighborhood for each filtered innovation $(\hat{\epsilon}_t, \hat{\zeta}_t)$. For any two innovations $(\hat{\epsilon}_t, \hat{\zeta}_t)$ and $(\hat{\epsilon}_s, \hat{\zeta}_s)$, define their distance as the Euclidean distance

$$d(t, s) = \sqrt{(\hat{\epsilon}_t - \hat{\epsilon}_s)^2 + (\hat{\zeta}_t - \hat{\zeta}_s)^2}.$$

For each given t , consider all distance $\{d(t, s) : s \neq t\}$ and rank them as $d_t^{(1)} \leq d_t^{(2)} \leq \dots \leq d_t^{(T-1)}$. For a given integer k , the k -distance of observation t is defined as $d_t^{(k)}$ and the k -distance neighborhood is defined as:

$$N_t^{(k)} = \left\{ s : d(t, s) \leq d_t^{(k)} \right\}.$$

Obviously, if all distances $d_t^{(j)}$ for $j = 1, 2, \dots, T-1$ are not equal, then the k -distance neighborhood $N_t^{(k)}$ contains exactly k points. They are the k points that are closest to the point t .

Next, we consider the reachability distance of s with respect to t . It is defined as:

$$\text{reach-dist}_k(s, t) = \max \left\{ d_t^{(k)}, d(s, t) \right\}.$$

In other words, for point s that is outside the k -distance neighborhood of t , $N_t^{(k)}$, the reachability distance to t is equivalent to the actual distance $d(s, t)$. However, for all points within the k -distance neighborhood of t , $N_t^{(k)}$, they are treated “equally” with the same reachability distance to t as the k -distance of t . Then, for each t we can define its local reachability density as

$$lrd_k(t) = \left(\frac{\sum_{s \in N_t^{(k)}} reach-dist_k(t, s)}{|N_t^{(k)}|} \right)^{-1}.$$

Intuitively, $1/lrd_k(t)$ is the average reachability distance of t with respect to each point in its own k -distance neighborhood. Therefore $lrd_k(t)$ would be high if such an average is low, i.e. when t is also likely to be within the k -distance neighborhood of s where s is in the k -distance neighborhood of t , i.e. t is also considered as a “neighbor” by its “neighbor”. By contrast, if $lrd_k(t)$ is low, t is considered to be far away from its “neighbor”. Finally, the local outlier factor of t is defined as

$$LOF_k(t) = \frac{\sum_{s \in N_t^{(k)}} \frac{lrd_k(s)}{lrd_k(t)}}{|N_t^{(k)}|},$$

, which is the ratio between the average local reachability density of the points in the k -distance neighbor of t and its own local reachability density. Intuitively, for a point that is deeply inside the cluster (i.e. not a local outlier), the LOF should be approximately 1 (see Breunig et al. (2000)). By ranking the LOF of all points in the scatters, one can identify the local outliers with the highest values of the LOF. A natural threshold is 1, but usually, a LOF value above 2 or 3 would be used for local outlier detection.

In this study, we use $k = 5$ and consider outliers with a LOF value above 3. The two scatters indicated by a red triangle in Figure 2 are the identified outliers. The highest LOF score is found for December 31, 2014, with a value at 3.36. The last business day of the year is traditionally one of the busiest in terms of number and turnover of transactions (see e.g. Heijmans and Heuver (2014) or Arciero et al. (2016)) due to balance sheet requirements from the bank supervisors or the shareholders (including activity for major clients of banks). Even though we correct for end-of-year effect in our analysis, this end of year seems to deviate from others. The second highest LOF score is found for April 1, 2010, with a value at 3.12. Like the end of year, there is a beginning of next quarter effect. Some transactions (lending, borrowing of liquidity) at the end of quarter are reversed at the beginning of the

next quarter. As we also include end-of-quarter effect, this suggest that this beginning of next quarter is different than others.

To further investigate the outlier days, we consider a new pair of indicators: HHI outgoing turnover (indicator 3) and net bilateral flows (NBF, indicator 10), and apply the same method to this pair, the result is shown in Figure 2b. Besides the two outliers found in Figure 2b, we also find two other outlier dates, i.e. January 30, 2012 and March 30, 2015. These dates are close to the end of the month and quarter of that year, respectively.

From these two pairs, we observe, that the relation between the HHI measure, irrespective of whether it is based on the outgoing or total turnover, and the net bilateral flows seem to formulate a triangular region in the scatter plot. That is, for a low value of HHI, the net bilateral flows indicator may spread from low to high, while for a high value of HHI, the net bilateral flows indicator must be at a low level. The lowest possible value of the net bilateral flows indicator seem to remain at a constant level as the value of HHI increases, while the highest possible value of the net bilateral flows indicator decreases as the value of HHI increases. The relation can also be switched: the lowest possible HHI value remains unchanged as the net bilateral flows increases, while the highest possible value of the HHI indicator decreases as the value of net bilateral flows indicator increases.

The two outliers discovered from the LOF method show different locations. The one on December 31, 2014, seems to lie at the upper edge of the triangular area. In other words, for the outlier date, given its net bilateral flows level, the HHI seems to be exceptionally high. By contrast, for the outlier detected on April 1, 2010, it seems that the HHI level is exceptionally low given its level of net bilateral flows. Nevertheless, for both dates, the absolute value of these two indicators are both at a moderate level. To further understand the mechanism behind such an outlier, we conduct a theoretical study regarding the relation between HHI and net bilateral flows in the next Section.

4.2 Multivariate outliers

To apply the LOF method for a policy analysis, we did an out-of-sample analysis aiming at detecting outliers in real time. This is similar to the univariate setup. The details of the out-of-sample is outlined as follows.

For each date t in the analysis period, for each indicator j , we fit the optimal model to the $n = 1250$ days proceeding date t . We obtain the estimated model coefficients as well as the filtered innovation

terms $\left\{\hat{\epsilon}_s^{(j)}\right\}_{s=t-n}^{t-1}$. Based on the the estimated model coefficients, and filtered innovation terms, we forecast each indicator at time t and obtain as the corresponding forecasting error, $\hat{\epsilon}_t^{(j)}$.

We collect the filtered innovations from the ten indicators on each day in the estimation period and the forecasting error on day t into a sample as $\left\{(\hat{\epsilon}_s^{(1)}, \dots, \hat{\epsilon}_s^{(10)})^T\right\}_{s=t-n}^t$. We treat them as an i.i.d. sample and calculate the LOF score of the last observation. In other words, we only examine whether the forecasting error on day t , as a 10-dimensional vector is a local outlier within the sample. This is comparable to the univariate out-of-sample, where we examine the tail probability of the forecasting error based on estimating the tail distribution function using the filtered innovations. With this analysis, we obtain the LOF score for each day in the out-of-sample testing period starting from Jan 1, 2015.

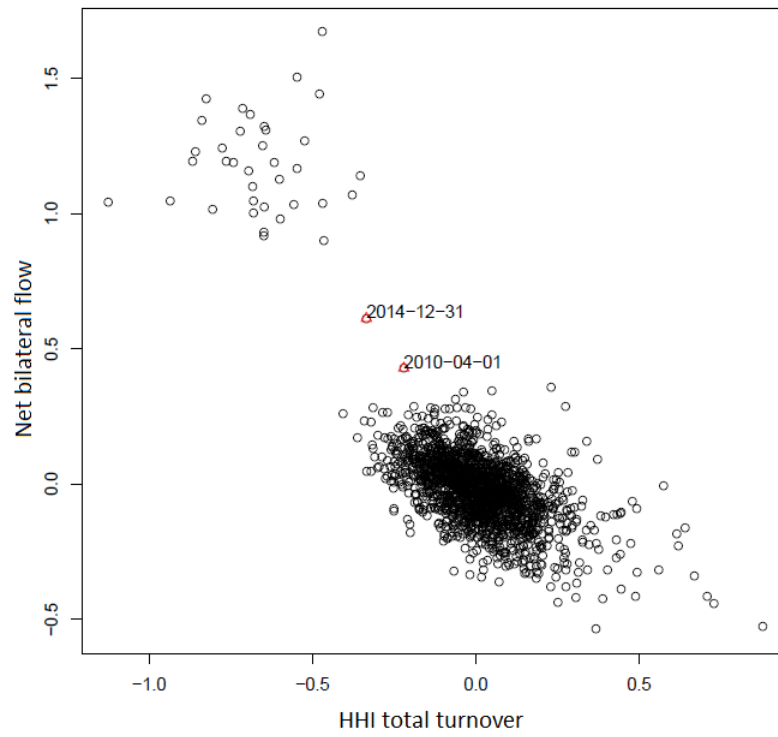
Table 5 shows the multivariate outliers using all indicators and indicators 2, 8 and 10 combined. First, we look at the outliers for the indicators 2, 8 and 10. These three indicators do not have univariate outliers, see section 3.3. As the table shows, our multivariate detection method detects three outlier dates. The first outlier date (May 25, 2015) has also been detected as an outlier date for three other indicators (3, 4 and 7). It appears that the individual indicators 2, 8 and 10 do not show an extreme value at each indicator separately, but they do show local outliers, when investigated with the multivariate outlier detection method. The two other dates neither have not been found before in the univariate analysis nor can they be linked to end or beginning of month. The multivariate outliers found using all indicators are all close to the end of the months.

Table 5: List of multivariate outliers for period 2015 and 2016.

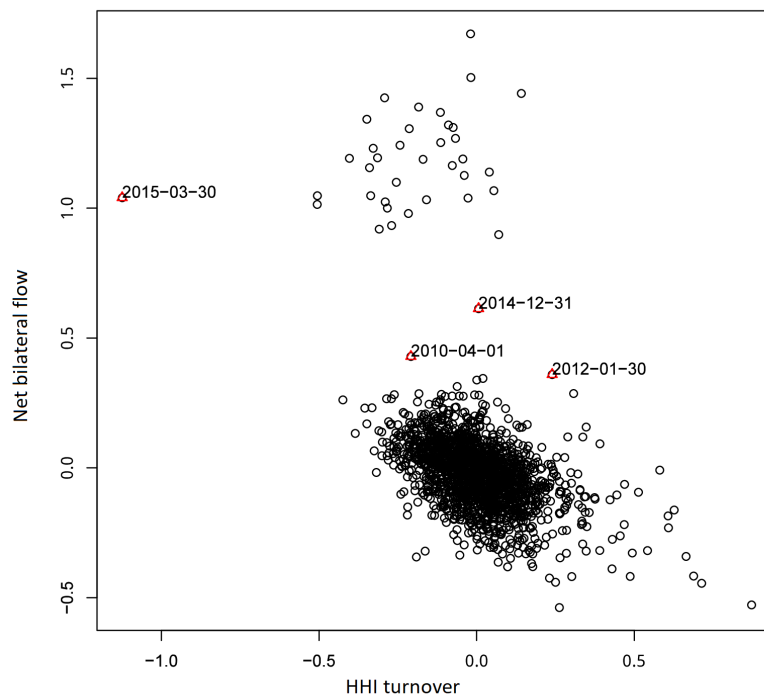
Date	LOF score	Comment
three indicators		
May 25, 2015	4.9	day after public holiday (Pentecost)
December 8, 2015	0.9	
January 21, 2016	2.5	
all indicators		
March 30, 2015	2.8	close to end of months
September 30, 2015	0.5	close to end of months
December 30, 2016	7.2	close to end of months

Figure 2: In-sample examples for local outlier detection.

(a) HHI total turnover vs NBF.



(b) HHI outgoing turnover vs NBF.



5 Theoretical model

In order to gain more insight regarding the bivariate outlier results obtained in section 4, we study the relation between the HHI and the net bilateral flows in a theoretical model. We consider a simple payment network consists of only three banks: A, B and C. Without loss of generality, we assume that the total sum of the payments made by all participants is one because both HHI turnover and the net bilateral flows are calculated as a ratio with respect to the total sum of the payments. Further, we assume that each bank pays as much as it receives, which means that the banks end with the same balance as they have started that day.¹⁰ This assumption helps to simplify the number of potential parameters in the network.

Figure 3 shows six theoretical examples of simplified payment networks of the payment system satisfying all assumptions. The corresponding adjacency matrix of each network with the liquidity flows is given below each network graph. Note that examples (a)-(e) involves gradually more payments in the system. Example (f) is the most general case following the assumptions. Examples (a)-(e) can be viewed as special cases of example (f) with specific parameter setups. Table 6 shows the results of the calculated HHI turnover and corresponding net bilateral flows for examples (a)-(e).

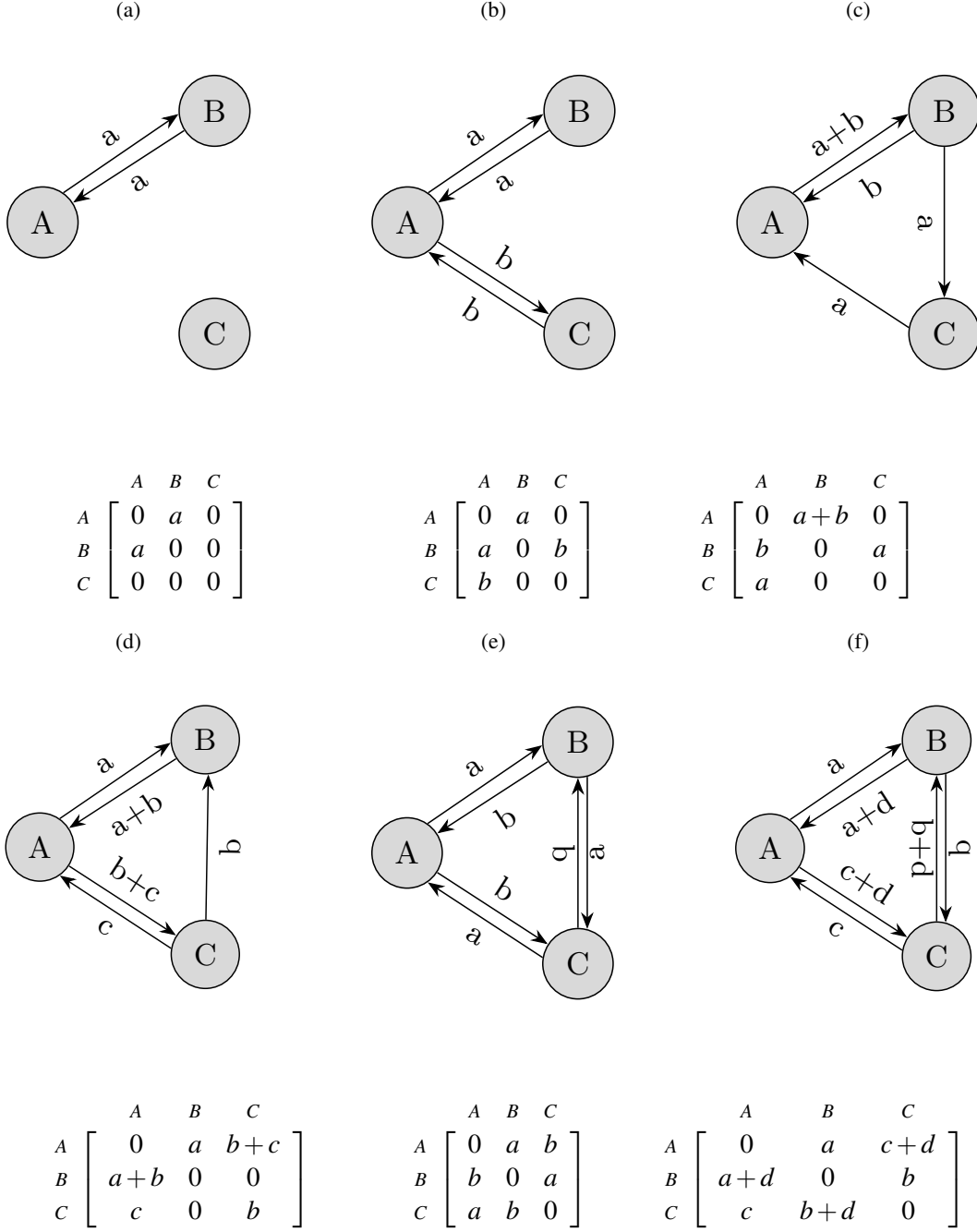
Table 6: Net bilateral flows and HHI total turnover for the first 5 theoretical payment networks.

graph nr	NBF	HHI	HHI min	HHI max	Conditions
a	0	$\frac{1}{2}$			$a = 1/2$
b	0	$\frac{1}{4} + a^2 + b^2$	$\frac{3}{8}$	$\frac{1}{2}$	$a + b = 1/2$
c	$3a$	$2(a + b)^2 + a^2$	$\frac{1}{3}$	$\frac{1}{2}$	$3a + 2b = 1$
d	$3b$	$(a + b + c)^2 + (a + b)^2 + (b + c)^2$	$\frac{1}{3}$	$\frac{1}{2}$	$2a + 3b + 2c = 1$
e	$3 a - b $	$\frac{1}{3}$			$a + b = 1/3$

We plot the potential paired values of (HHI, NBF) stemming from these five examples in Figure 4. The cases (a) (b) (c) and (e) are illustrated by corresponding curves. For the case (d), it is handled in the following Lemma.

¹⁰Cash managers of banks have to manage the liquidity on their account. Due to reserve maintenance requirements banks do need to have a certain pre defined amount of liquidity during the reserver maintenance period, which is typically 4 to 7 weeks in the Eurosystem. Although the actual liquidity at the end of day may vary a bit, banks do manage their liquidity actively and do normally not allow very large deviations in it.

Figure 3: Simplified examples of payment networks with their correspondind adjancency matrix.



Lemma 1 For the network (d), for any given $0 \leq b \leq 1/3$, we have that

$$\frac{3-2b+3b^2}{8} \leq HHI(b) \leq \frac{1-2b+3b^2}{2}.$$

Notice that the upper bound of $HHI(b)$ corresponds to $a = 0$ or $c = 0$. When reaching this upper bound, the network simplifies to the case (c). For the lower bound, we plot the curve (HHI, NBF) in

Figure 4.

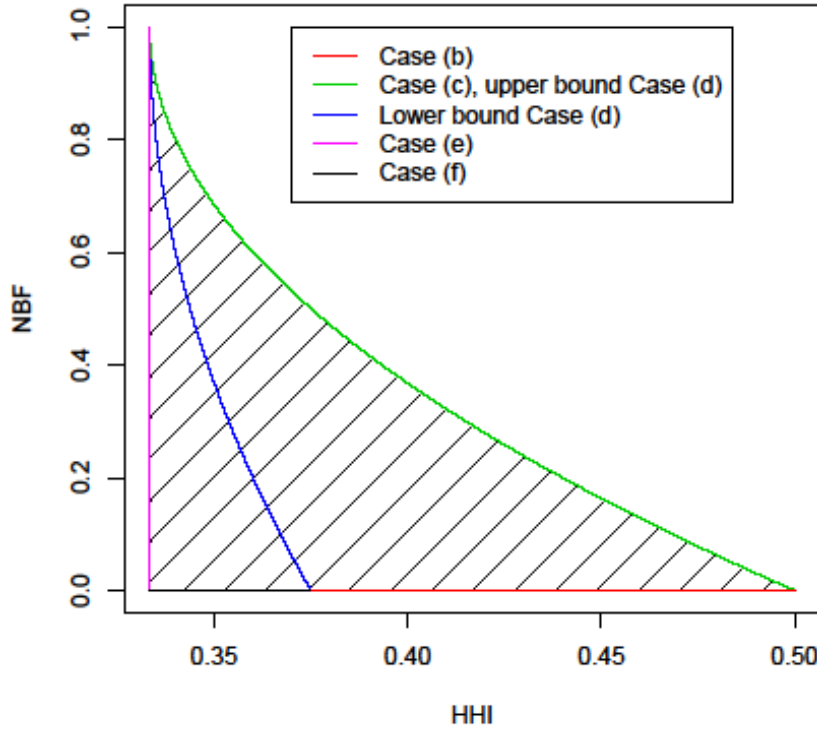
Lemma 2 For the network (f), for any given $0 \leq d \leq 1/3$, we have that

$$\frac{1}{3} \leq HHI(d) \leq \frac{1 - 2d + 3d^2}{2}.$$

Notice that the upper bound again corresponds to the case (c). The lower bound corresponds to the case (e).

Overall, from the analysis for the six networks, we can draw the following conclusions. Firstly, the potential pair of (HHI, NBF) is bounded in the pseudo-triangle area indicated by the shaded area in Figure 4. The area is a pseudo-triangle, because one edge is defined by a curve (case (c)). Notice that the scatter plots in Figure 2 resembles such a pseudo-triangle.

Figure 4: Potential values for HHI total turnover and net bilateral flow.



Secondly, in the upper side of the pseudo-triangle, NBF is at a high level indicating that there are “circular” payments in the system which cannot be netted out at a bilateral level. These are situations where external liquidity is needed to get all payment settled, even if no liquidity is needed for any bank in the system, because all in-flows and out-flows are assumed to be balanced. Notice that the scatter plot in Figure 2 shows that there are two different scenarios in the data. Most of the scatters

resemble the low NBF high HHI case. This is the case where bilateral netting can be conducted (case (b)). A few scatters in the left-upper corner of the pseudo-triangle resembles exceptional days where large circular payments are conducted. The identified outlier dates are in between these two scenarios. Thirdly, a pair (HHI, NBF) in the left-bottom area of the pseudo-triangle, i.e. below the lower bound defined in case (d), can be realized if and only if payments in both directions are executed between any two banks. In other words, the payment activities are intensive. For any given level of NBF, the highest possible HHI is reached at case (c) where a large circular payment involving all banks is observed. For the fixed level of NBF, as the HHI level goes down, payments in the reverse direction of the large circular payment are observed (case (d)). When the level goes below the lower bound in case (d), a reverse circular payment definitely exists (case (f)).

Although our theoretical model is only based on three banks, we can extend this result to a larger payment network. For a given NBF level, high HHI corresponds to a large but long circular payments involving many banks in the system. Low HHI corresponds to multiple circular payments, such as a reverse circular payment, with uneven size affecting different part of the system. Both cases need external liquidity to settle the payments, but the operation in the low HHI case is even more complex. Now, we go back to the outliers detected in Figure 2 and interpret their unusualness. The identified local outlier on December 31, 2014 lies close to the upper bound of the HHI given its NBF level. It shows that a large circular payment exists on this day. The identified local outlier on April 1, 2010 lies close to the lower bound of the HHI given its NBF level. This indicates that the payment networks on April 1, 2010 may suffer from a complex circular payment structure involving multiple circular payments as in case (f) of the theoretical model.

6 Policy analysis and summary

In this paper we showed a univariate and a bivariate method for detecting outliers in TARGET2 risk indicator. The univariate analysis using extreme value theory is able to detect extreme values of an indicator time series. The bivariate (multivariate with two indicators) allows for detecting outliers that are not necessarily extreme values, but also values that are not likely seen between a pair of indicators. These so called local outliers can be detected using the local outlier factor method. Similarly, local outliers can be detected using multiple indicators.

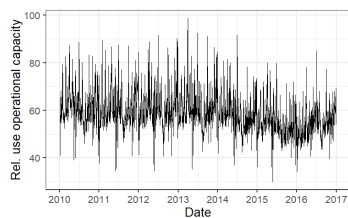
Many outliers detected with the aforementioned methods are close to the end or near the beginning of a month, quarter or year, even though our indicator time series have been corrected for end of month (quarter and year) effects. It seems that at those outlier dates, banks behave somewhat different than usual. However, we have also found outlier dates we could not connect to a certain event. These outlier dates can be further investigated by the expert in charge to see what the cause of this different behavior could be.

Appendix A Risk indicator time series

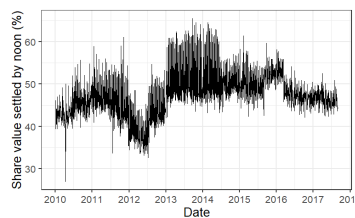
Figures 5a to 5i show the time series of the risk indicators used in this paper. The number in the figure caption refers to the number assigned in Table 2.

Figure 5: Time series of the risk indicators of Table 2. The numbers refer to the numbers of the risk indicator in that table.

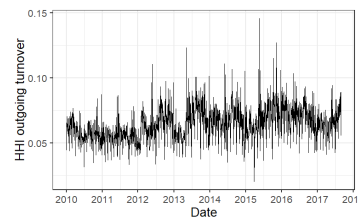
(a) Use of operational capacity (indicator 1).



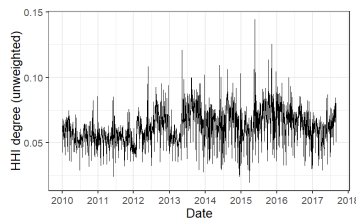
(b) Throughput guideline (indicator 2).



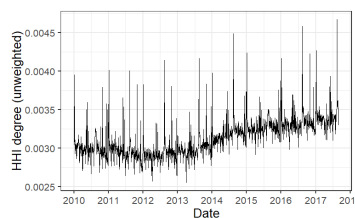
(c) HHI outgoing turnover (indicator 3).



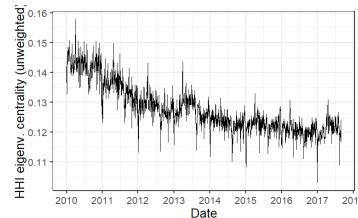
(d) HHI total turnover (indicator 4).



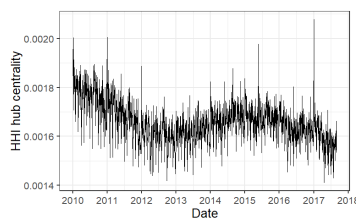
(e) HHI degree (indicator 5).



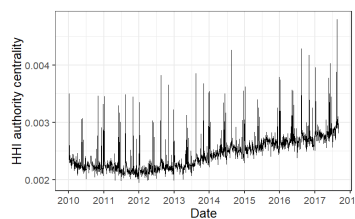
(f) HHI eigenvector centrality (indicator 6).



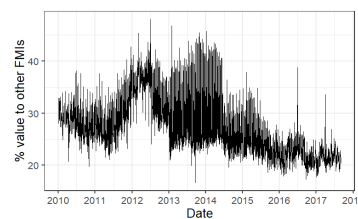
(g) Hub centrality (indicator 7).



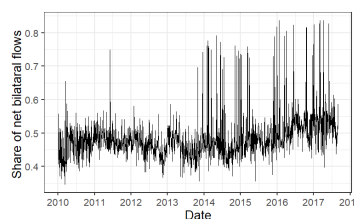
(h) Authority centrality (indicator 8).



(i) Turnover to other FMIs (indicator 9).



(j) Net bilateral flows (indicator 10).



References

- Abbink, K., Bosman, R., Heijmans, R., and van Winden, F. (2017). Disruptions in Large Value Payment Systems: An Experimental Approach. *International Journal of Central Banking*, 13(4):63–95.
- Arciero, L., Heijmans, R., Heuver, R., Massarenti, M., Picillo, C., and Vacirca, F. (2016). How to Measure the Unsecured Money Market? The Eurosystem’s Implementation and Validation using TARGET2 Data. *International Journal of Central Banking*, March:247–280.
- Baek, S., Soramäki, K., and Yoon, J. (2014). Network indicators for monitoring intraday liquidity in bok-wire+. *Journal of Financial Market Infrastructures*, 2(3):37–66.
- Bech, M. and Garratt, R. (2003). The Intraday Liquidity Management Game. *Journal of Economic Theory*, 109:198–210.
- Berndsen, R. and Heijmans, R. (2017). Risk indicators for financial market infrastructures: From high frequency transaction data to a traffic light signal. *DNB Working Paper*, 557.
- Breunig, M. M., Kriegel, H.-P., Ng, R. T., and Sander, J. (2000). LOF: identifying density-based local outliers. In *ACM Sigmod Record*, volume 29, pages 93–104. ACM.
- CPSS (2012). Principles for Financial Market Infrastructures: Disclosure Framework and Assessment Methodology. *Bank for International Settlements*.
- de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer Science & Business Media.
- Diehl, M. (2013). Measuring Free Riding in Large-Value Payment Systems: The Case of TARGET2. *Journal of Financial Market Infrastructures*, 1(3):31–53.
- Embrechts, P., Mikosch, T., and Klüppelberg, C. (1997). *Modelling extremal events: for insurance and finance*. Springer-Verlag.
- Heemeijer, P. and Heijmans, R. (2015). Central bank intervention in large value payment systems: An experimental approach. *Journal of Financial Market Infrastructures*, 3(3):17–49.
- Heijmans, R. and Heuver, R. (2014). Is this Bank Ill? The Diagnosis of Doctor TARGET2. *Journal of Financial Market Infrastructures*, 2(3):3–36.
- Heijmans, R., Heuver, R., Lelyveld, I., and Levallois, C. (2016). Dynamic visualization of large financial networks. *Journal of Network Theory*, 2(2):1–23.
- Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174.
- Kaliontzoglou, A. and Müller, A. (2015). Implementation aspects of indicators related to payments timing. In Diehl, M., Alexandrova-Kabadjova, B., Heuver, R., and Martinez-Jaramillo, S., editors, *Analyzing the Economics of Financial Market Infrastructures*, pages 169–190.
- Li, F. and Pérez-Saiz, H. (2016). Measuring systemic risk across financial market infrastructures. *Bank of Canada Staff Working Paper*, March.
- Loureiro, A., Torgo, L., and Soares, C. (2004). Outlier detection using clustering methods: a data cleaning application. In *Proceedings of KDDNet Symposium on Knowledge-based systems for the Public Sector*. Springer.

- Massarenti, M., Petriconi, S., and Lindner, J. (2012). Intraday Patterns and Timing of TARGET2 Interbank Payments. *Journal of Financial Market Infrastructures*, 1(2):3–24.
- Roberts, S. J. (1999). Novelty detection using extreme value statistics. *IEE Proceedings-Vision, Image and Signal Processing*, 146(3):124–129.
- Roberts, S. J. (2000). Extreme value statistics for novelty detection in biomedical data processing. *IEE Proceedings-Science, Measurement and Technology*, 147(6):363–367.
- Sohn, H., Allen, D. W., Worden, K., and Farrar, C. R. (2005). Structural damage classification using extreme value statistics. *Journal of Dynamic Systems, Measurement, and Control*, 127(1):125–132.
- Squartini, T., van Lelyveld, I., and Garlaschelli, D. (2013). Early-warning signals of topological collapse in interbank networks. *Scientific Reports*, 3(3357):1–9.
- Sundaram, S., Strachan, I. G., Clifton, D. A., Tarassenko, L., and King, S. (2009). Aircraft engine health monitoring using density modelling and extreme value statistics. In *Proceedings of the 6th International Conference on Condition Monitoring and Machine Failure Prevention Technologies*.
- Timmermans, M., Heijmans, R., and Daniels, H. (2018). Cyclical patterns in risk indicators based on financial market infrastructure transaction data. *Quantitative Finance and Economics*, 2(3):615–636.
- Triepels, R., Daniels, H., and Heijmans, R. (2018). *Detection and Explanation of Anomalous Payment Behaviour in Real-Time Gross Settlement Systems*, volume 321. Springer Cham.
- Van Ark, T. and Heijmans, R. (2016). Granularity, a blessing in disguise: Transaction cycles within real-time gross settlement systems. *Journal of Financial Market Infrastructures*, 5(2):29–52.
- Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815.

Previous DNB Working Papers in 2019

- No. 622 **David-Jan Jansen**, Did Spillovers From Europe Indeed Contribute to the 2010 U.S. Flash Crash?
- No. 623 **Wilko Bolt, Kostas Mavromatis and Sweder van Wijnbergen**, The Global Macroeconomics of a trade war: the EAGLE model on the US-China trade conflict

DeNederlandscheBank

EUROSYSTEEM

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl