



The concept of Tax Gaps

Report III: MTIC Fraud Gap estimation methodologies



EUROPEAN COMMISSION

Directorate-General for Taxation and Customs Union
Directorate C — Indirect taxation and tax administration
Unit C.4 — Tax administration and fight against tax fraud

Contact: TAXUD UNIT C4

E-mail: TAXUD-UNIT-C4@ec.europa.eu

European Commission
B-1049 Brussels

The concept of Tax Gaps

Report III: MTIC Fraud Gap estimation methodologies



***Europe Direct is a service to help you find answers
to your questions about the European Union.***

Freephone number (*):

00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

LEGAL NOTICE

This document has been prepared by the participants of the Tax Gap Project Group (TGPG-VF). The views expressed in this report are those of the TGPG members and do not necessarily represent the views of the participant's national administration or those of the European Commission. The Commission cannot be held responsible for any use which may be made of the information contained therein.

More information on the European Union is available on the Internet (<http://www.europa.eu>).

Luxembourg: Publications Office of the European Union, 2018

ISBN 978-92-79-97236-2
doi: 10.2778/418684

© European Union, 2018
Reproduction is authorised provided the source is acknowledged.

TABLE OF CONTENT

1.	INTRODUCTION	5
2.	VAT FRAUD SCHEMES	7
2.1.	Definition of VAT fraud	7
2.1.1.	The VAT system	7
2.1.2.	VAT fraud terminology	7
2.2.	Classical VAT fraud	8
2.2.1.	Underreported sales	8
2.2.2.	Failure to register.....	8
2.2.3.	Omission of self-deliveries.....	8
2.2.4.	VAT collected but not remitted.....	8
2.2.5.	False claims for credit or refund	8
2.2.6.	Credit claimed for VAT on purchases that are not creditable	8
2.3.	Organized VAT fraud	9
2.3.1.	VAT rules for intra-Community (IC) transactions	9
2.3.2.	Missing trader intra-Community acquisition fraud (ICA fraud) ..	10
2.3.3.	Missing trader intra-Community (MTIC) fraud	11
2.3.4.	Missing trader intra-Community (MTIC) fraud – Cross - invoicer.....	12
2.3.5.	Missing trader intra community (MTIC) fraud – Contra-trading.	12
2.3.6.	Triangular VAT fraud.....	13
2.3.7.	Domestic sales designed as intra-Community supplies	14
3.	MTIC IN LITERATURE	15
3.1.	General literature about MTIC fraud	15
3.2.	Literature about the estimation of VAT fraud	16
4.	MTIC FRAUD ESTIMATION METHODS	19
4.1.	Introduction.....	19
4.2.	Top-down approaches	19
4.2.1.	Intrastat mirror statistics	20
4.2.2.	Method based on the variation of macro variables	21
4.3.	Bottom-up approaches	22
4.3.1.	Measuring MTIC fraud based on risk analysis (Belgium)	23
4.3.1.1.	Profiling ICA Cross-Checking and the predictive modelling.....	23
4.3.1.2.	The Belgian Experiment	24
4.3.2.	Accrued tax by risky people (Bulgaria)	25
4.3.2.1.	Methodology – selecting the risky persons –risk register	25
4.3.2.2.	Measurement of the tax gap from VAT frauds: Sales method ...	26
4.3.2.3.	Sources of information.....	27
4.3.3.	Measuring MTIC fraud based on risk analysis (Finland)	28
4.3.3.1.	Selection or Profiling method	28
4.3.3.2.	Estimation method	29
4.3.4.	Tax gap extrapolation from random audit data.....	31
4.3.4.1.	Data collection	31
4.3.4.2.	Extrapolation	32
4.3.4.3.	Selection random samples for MTIC estimation and stratification	32

4.3.4.4. Advantages and disadvantages	32
4.3.5. K-means clustering	33
4.4. Comparison of the different methods with regards to possible utilities	34
5. DATA SOURCES	35
5.1. National VAT declarations	35
5.2. Recapitulative statements and VIES data	35
5.3. Intrastat data.....	35
5.4. Audit findings.....	36
5.5. Other information	36
5.6. Overview of data used by the various methods	37
6. A EUROPEAN APPROACH.....	38
6.1. The TNA Project	38
6.1.1. Context.....	38
6.1.2. Evaluation of MTIC Gap.....	39
7. CONCLUSIONS AND RECOMMENDATIONS.....	40
7.1. Conclusions	40
7.2. Recommendations	40
REFERENCES	42
ANNEX I. - LIST OF PARTICIPANTS	45
ANNEX II. - LIST OF PRESENTATIONS.....	46

Index of Figures

Figure 1: Relationship between VAT gap, VAT fraud and MTIC fraud	7
Figure 2: MTIC fraud - Acquisition fraud	10
Figure 3: MTIC fraud – Carousel fraud	11
Figure 4: MTIC fraud - Carousel fraud including a buffer company	12
Figure 5: MTIC fraud – Cross invoicer	12
Figure 6: MTIC fraud – Contra trading.....	13
Figure 7: MTIC fraud – Triangular VAT fraud.....	13
Figure 8: Fictitious intra-Community supplies	14
Figure 9: Cases with positive ICA discrepancy	24
Figure 10: Total of positive ICA discrepancies	25
Figure 11: Principle scheme of the methodology	26
Figure 12: The result of profiling	29
Figure 13: Cut distributions	29
Figure 14: K-means clustering	33

Index of Tables

Table 1: The estimation of the number of F-ICAs in the population (X).....	30
Table 2: The estimated number of fraudulent intra-Community acquirers	30
Table 3: Comparison of the different models	34
Table 4: Overview of data used by the various methods	37

Abbreviations and ACRONYMS

ATP	Arbitrage Pricing Theory
BE	Belgium
BC	Buffer Company
BG	Bulgaria
BR	Broker Company
CASE	Centre for Social and Economic Research
CC	Conduit Company
CIT	Corporate Income Tax
CZ	Czech Republic
DS	Domestic Supply
DG TAXUD	Directorate-General for Taxation and Customs Union
ELO	Eurofisc Liaison Officer
ES	Spain
ESTAT	Eurostat
ETCB	Estonian Tax and Customs Board
EC	European Commission
EU	European Union
FI	Finland
F-ICA	Fraudulent Intra-Community Acquirer
GDP	Gross Domestic Product
GNI	Gross National Income
HMRC	Her Majesty Revenue and Customs
HU	Hungary
IC	Intra-Community
ICA	Intra-Community Acquisition
ICS	Intra-Community Supply
IT	Italy
MoF	Ministry of Finance
MS	Member State
MT	Missing Trader
MTIC	Missing Trader Intra Community
LT	Lithuania
NL	Netherlands
NRA	National Revenue Agency
PL	Poland
PT	Portugal
RO	Romania
SK	Slovakia
TGPG-VF	Tax Gap Project Group on methodologies to estimate MTIC fraud
UK	United Kingdom
VAT	Value Added Tax
VAT Gap	Difference between the VAT revenues theoretically established by legislation and actual collections
VIIES	VAT Information Exchange System

1. INTRODUCTION

Value Added Tax (VAT) is an important and often main contributor to the State Budget of Member States. Collecting the correct amount of VAT is an ongoing challenge for national tax authorities as to secure VAT revenues, to ensure fair competition between businesses and to avoid economic distortions of the national and internal market.

The 2018 VAT Gap¹ study indicates a major gap of EUR 147 billion in 2016 between the VAT that should have been paid and the VAT actually collected.² There are different reasons for this lack of compliance, ranging from negligence, omissions, non-deliberate errors, differences in interpretation, lack of knowledge and insolvencies to deliberate actions such as tax fraud, tax evasion and tax avoidance. However, the current VAT Gap study does not allow for a breakdown into the different causes of the gap.³

Often the question is raised which part of the VAT gap can be linked to VAT fraud and evasion. This call coincides well with the political priorities of the current Juncker Commission, where in his fourth priority President Juncker stresses the need for "A deeper and fairer internal market with a strengthened industrial base". He highlights that "We need more fairness in our internal market. We should step up our efforts to combat tax evasion and tax fraud, so that all contribute their fair share". In this framework, the problem of VAT fraud receives high attention in the Communications on an Action plan on VAT⁴ and in the recent VAT proposal for a VAT definitive regime.⁵

The challenge for the Fiscalis Project Group on Tax Gap – subgroup VAT fraud estimation methodologies (TGPG-VF) is how to identify the part of the VAT gap related to VAT fraud and evasion. As VAT fraud and evasion can be committed in various ways, it is important to limit the scope of the exercise. Although other types of VAT fraud and evasion are important too, for several reasons it was decided to focus on Missing Trader Intra Community (MTIC) fraud. MTIC fraud is a severe type of fraud, concentrated and organised. It has a European Union (EU) dimension since it requires cross border trade and all Member States are affected to a greater or lesser extent. Furthermore, there is experience with MTIC fraud schemes and more data seems to be available than for other types of fraud.

A short survey was sent to all Member States to enquire about their experiences with estimating VAT fraud and/or MTIC fraud. The results showed that there is little experience in estimating the size of these types of fraud. Only a few Member States and some external parties (academics, tax consultancy firms, and international organisations) have calculated the size of MTIC fraud at national level or European level.⁶ According to these calculations the size differs significantly ranging from EUR 20 billion up to more than EUR 100 billion a year. Therefore, there is a need to explore in more detail the

¹ VAT Gap = the difference between the VAT revenues theoretically established by legislation and actual collections.

² CASE c.s., *Study and Reports on the VAT Gap in the EU-28 Member States: 2018 Final Report*, TAXUD/2015/CC/131, Warsaw, 11 September 2018.

³ It is not in the terms of reference of the contractor to calculate the size of VAT fraud and there is not sufficient data available at EU level for a breakdown into the different reasons of the gap.

⁴ Communication on an *Action Plan on VAT, towards a single EU VAT area - Time to decide*, COM(2016) 148 final, Brussels, 7.4.2016 and Communication on *the follow-up to the Action Plan on VAT - Towards a single EU VAT area - Time to act*, COM(2017) 566 final, Brussels, 4.10.2017

⁵ Proposal for a Council Directive amending Directive 2006/112/EC as regards harmonising and simplifying certain rules in the value added tax system and introducing the definitive system for the taxation of trade between Member States, COM(2017) 569 final, Brussels, 4.10.2017

⁶ Examples: The study "*Implementing the 'destination principle' to intra-EU B2B supplies of goods*", final report, 30 June 2015, mentions a figure of EUR 50 billion as VAT fraud. Europol has estimated the scale of VAT fraud on EUR 100 billion. Some academics have made calculations of the size of VAT fraud, ranging from EUR 20 - EUR 35 billion to EUR 94 billion a year.

existing methods and whether it would be possible to design a methodology that could estimate the size of MTIC fraud at national and European level. For the latter an EU wide method would be helpful for reasons of comparison in time and between Member States. In addition, it may contribute to assessing the efficiency of tax collection as well as the design of specific anti-VAT fraud measures.

This report describes existing methodologies that are used to estimate the size of MTIC fraud and explores whether these methods can be adapted to a European approach. It should be emphasised that this report does not include calculations of MTIC fraud in the EU.

The report is the third one in a row.⁷ It is prepared by the TGPG-VF. This group was established under the Fiscalis 2020 programme as a continuation of the Fiscalis Tax Gap Project Group (FPG/041). The European Commission, Directorate-General for Taxation and Customs Union (DG TAXUD) coordinate its work.⁸

The report is based on the contributions of the TGPG-VF participants. To estimate VAT fraud it is necessary to understand the mechanism of VAT fraud. Therefore, the second chapter of the report contains descriptions of VAT fraud types and more specifically of MTIC fraud. The third chapter provides a (non-limited) overview of existing scientific literature (articles, books, etc.) on MTIC fraud and methodologies to estimate this fraud. The fourth chapter describes VAT fraud estimation methodologies and is divided into two parts with different approaches: top-down and bottom-up methods. Data sources that could be useful for the estimation of VAT fraud are presented in chapter five. Chapter six reflects on the possibility of a European approach. Finally, in chapter seven the report ends with conclusions and recommendations.

The contributions for this report were provided by Mr Yannic Hulot (BE), Mr Jíří Vlach (CZ), Ms Estrella Gonzalez and Mr Ramiro Ledo (ES), Mr Aki Savolainen and Ms Kirsi Ristola (FI), Mr Ván Bálint (HU), Ms Elena D' Agosto (IT), Ms Agné Jakubauskaitė (LT), Mr Lou Verberkt and Mr Menno Griffioen (NL), Mr Tomasz Mazur and Ms Teresa Chomiczewska (PL), Ms Carla Pereira Rodrigues (PT), Mr Victor Ogneru (RO), Ms Martina Motkova (SK), Mr Tanel Puetsep (EC) and Ms Jannetje Bussink (EC, Chair TGPG-VF). For a more detailed list of participants of the TGPG-VF, see Annex I. The contributions were compiled and the report was edited by DG TAXUD.

The TGPG-VF held seven meetings. At the meetings, both guest speakers and members of the project group gave several presentations. The TGPG-VF is especially grateful for the stimulating presentations and useful inputs of Mr Grzegorz Poniatoski (CASE), Dr Marius Cristian Frunza (Schwarzthal Kapital), Ms Iona Stretton and Mr Kishel Patel (Ernst & Young), Ms Kirsi Ristola (Finnish Tax Administration), Mr Nikolay Petkov (Bulgarian National Revenue Agency), Mr Ádám Fajkus (Hungarian Tax and Customs Administration) and Ms Catrine Boogh-Dahlberg (Eurostat). For a more detailed list of presentations, see Annex II.

The meetings of the TGPG-VF were hosted by the Dutch Tax and Customs Administration, the Portuguese Tax and Customs Authority and the European Commission. The TGPG-VF is thankful to these organisations for making their facilities available.

⁷ Previous reports are "The Concept of Tax Gaps – Report on VAT Gap Estimations" and "The Concept of Tax Gaps - Report II: Corporate Income Tax Gap Estimation Methodologies".

⁸ Estimating a tax gap is difficult both in terms of methodological consideration and data needs. To pool knowledge and share experience in estimating the tax gap, the Tax Gap Project Group (TGPG) was established under the Fiscalis 2020 Programme. The group initially focused on the VAT gap, which led to the publication of a report on "The Concept of Tax Gaps – Report on VAT Gap Estimations". Subsequently, the TGPG decided to split into two subgroups, one focusing on direct tax gap methodologies, particularly methodologies for corporate income tax gaps (CIT gap), and the other one on VAT fraud gap estimation methodologies, with a special focus on missing trade intra-Community fraud (MTIC fraud).

2. VAT fraud schemes

The current VAT system is vulnerable to fraud and evasion. VAT fraud can be committed in various ways ranging from a simple under-declaration of sales to sophisticated cross border fraud schemes. This chapter will briefly highlight two groups of VAT fraud schemes, i.e. so-called classical VAT fraud and organized VAT fraud.

2.1. Definition of VAT fraud

2.1.1. The VAT system

VAT is a general tax that in principle applies to all commercial activities involving the production and distribution of goods and the provision of services. It is a consumption tax, borne ultimately by the final consumer. VAT is collected fractionally via a system of partial payments whereby taxable persons (i.e. VAT-registered businesses) deduct from the VAT they have collected ("output VAT") the amount of tax they have paid to other taxable persons on purchases for their business activities ("input VAT"). Any excess credits this creates are refunded to the taxpayer. VAT is an indirect tax since the revenues are paid to the tax authorities by the seller of the goods and services, instead of the buyer who actually pays the VAT. This VAT mechanism is susceptible to several types of fraud.

2.1.2. VAT fraud terminology

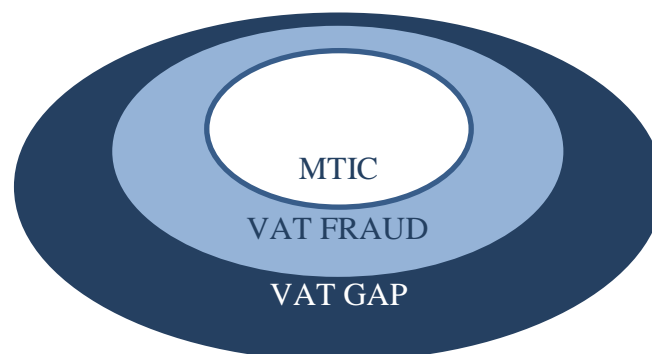
There are various definitions of the terms "VAT fraud and evasion". In this report these terms are defined as follows:

VAT evasion generally comprises illegal arrangements where tax liability is hidden or ignored, i.e. the taxpayer pays less tax than he/she is supposed to pay under the law by hiding income or information from the tax authorities.

VAT fraud is a form of deliberate evasion of tax which is generally punishable under criminal law. The term includes situations in which deliberately false statements are submitted or fake documents are produced. It is organised fraud and includes national and cross border transactions.

MTIC fraud is a specific form of VAT fraud. VAT is stolen from a government by organised criminal activity, which exploits cross border trading where the movement of goods between jurisdictions is VAT-free. This allows the fraudster (the person who commits fraud) to charge VAT on the sale of goods, and then instead of paying this to the government's collection authority, simply disappear, taking the VAT with him.

Figure 1: Relationship between VAT gap, VAT fraud and MTIC fraud⁹



⁹ The picture does not represent the relative size of the VAT fraud and MTIC fraud, but is meant as an illustration of the relationship between the three terms used in this report.

For the purpose of this report, it is important to stress that the focus will be on MTIC fraud as a result of actions by more than one person.

2.2. Classical VAT fraud

The report distinguishes between classical VAT fraud and organized VAT fraud. The term "classical VAT fraud" can be characterized as "VAT evasion". Examples of this type of VAT fraud are listed in the following subparagraphs for information.

2.2.1. Underreported sales

In this type of fraud the trader reports only a proportion of the sales, falsifying records and accounts to match, or may even make some sales 'off the books' entirely. The loss of VAT revenue arises from customers registered for VAT seeking the correspondent refund, the trader not collecting tax or the trader generating excess credits to be refunded. This type of VAT fraud is more common in retail trade and service sectors, where VAT deductible on input is small relative to VAT taxable on output.

2.2.2. Failure to register

The trader fails to register for two reasons: to avoid paying tax and to avoid the VAT compliance costs. Failure to register is more common among small businesses operating close to the level of turnover at which registration becomes compulsory or at which tax is due.

2.2.3. Omission of self-deliveries

This type of evasion is also more common among small businesses. The business output is consumed either by the business owner or by the employees, and not declared. The impact on revenue is usually rather low.

2.2.4. VAT collected but not remitted

There is another VAT evasion opportunity that arises in the supply chain transactions and is referred to as VAT not remitted. Usually, sellers charge VAT on their sales and show the amount of tax liability in the invoices sent to the purchasers. With this action the seller is entitled to receive VAT from the customer. At this stage, sellers collect VAT but they fail to remit the VAT to the State budget.

2.2.5. False claims for credit or refund

This type of fraud involves the setting up of false enterprises and their registration for tax purposes. These enterprises then produce invoices for non-existing purchases, which give grounds for VAT refund claims. The purchase might even be real, but exaggerated. This scheme is also connected to fake exports.

2.2.6. Credit claimed for VAT on purchases that are not creditable

Two types of this kind of fraud can be identified. First, when traders supply a variety of outputs, some subject to VAT and others exempt, they have an incentive to allocate inputs to production of the taxed items (in respect of which input tax credit is available) rather than the exempt (for which it is not). Second, items bought for private consumption may be misrepresented as business inputs, allowing the VAT to be recovered (and income tax liability reduced).

2.3. Organized VAT fraud

The term "organized VAT fraud" is similar to the term "VAT fraud". This section describes the main schemes of organized VAT fraud with cross border transactions.¹⁰ Although the different schemes may apply to all international transactions, the focus will be on those schemes showing intra-Community (IC) transactions, in other words: **MTIC fraud**.

In the overview of VAT fraud schemes it is indicated which Member State is bearing the VAT loss. However, it should be mentioned that due to the complexity of the schemes it is almost impossible to identify clearly which Member State(s) and to what extent they are confronted with VAT losses.

2.3.1. VAT rules for intra-Community (IC) transactions

In case of trade among EU Member States, the general VAT rules indicate that VAT-registered suppliers are entitled to apply a zero rate on their intra-Community sales (ICS) to VAT-registered buyers in other Member States under the condition that the goods are dispatched or transported from one EU Member State to a destination in another Member State.¹¹ VAT-registered IC suppliers are able to claim VAT on their purchases, often resulting in a request for a VAT refund.

The purchaser, not paying VAT to the supplier, must declare an intra-Community acquisition (ICA) in the Member State of destination. This VAT is deductible. Then, VAT becomes chargeable on the subsequent sales in the Member State of destination.

The taxation mechanism of IC transactions offers fraudsters an outstanding opportunity to steal money from the State. Starting from simple fraud schemes fraudsters designed more and more sophisticated structures to raise more revenues without being detected. In this section the main IC VAT fraud schemes are presented.

Before presenting these schemes the role of the different parties involved in the IC frauds is examined.

- The Missing Trader (MT) is a (fictitious) company employed to simulate an IC transaction in order to obtain a fraudulent gain from VAT charged on a subsequent transaction. When it is time to pay the VAT to the tax authorities the MT disappears.
- The Buffer Company (BC) is an entity that acts as a normal trader, purchasing and supplying goods and/or services on the domestic market. It is employed in the fraud chain after the MT to hide the fraud scheme and make investigations more difficult.
- The Broker (BR) is the final link of the fraudulent scheme and is situated in the same Member State as the MT. It purchases goods and/or services from the BC and supplies them to an operator established in the domestic market or in another Member State. In case it makes an IC supply, it will claim for the refund of the VAT paid on its purchases.
- The Conduit Company (CC) is a trader that participates in a transaction that is connected with the fraudulent evasion of VAT in another Member State.

¹⁰ The descriptions follow the document "Good Practice Guide to tackling Intra-Community VAT fraud", updated version 2013, Fiscalis Project Group – this guide is for internal use by Member States only.

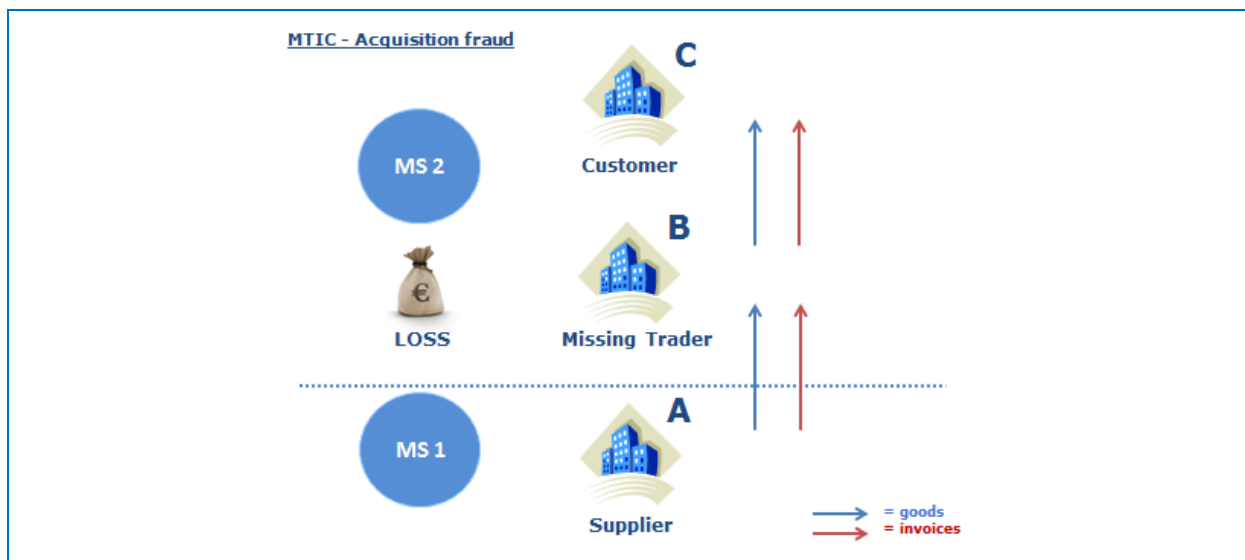
¹¹ The goods must physically leave the Member State of the supplier.

2.3.2. Missing trader intra-Community acquisition fraud (ICA fraud)

This scheme, called "acquisition fraud", shows the basic conditions for an ICA fraud. It requires two Member States and three parties: one supplier in Member State 1 (MS 1) and one trader/buyer plus one consumer in Member State 2 (MS 2).

Company A (supplier) in MS 1 sells goods to company B (buyer) in MS 2 applying the zero rate on the intra-Community supply (ICS). Company B in MS 2 has to declare an intra-Community acquisition (ICA) which is deductible, as he sells the goods onwards to his customers in MS 2 charging the corresponding VAT (i.e. so-called domestic supply (DS)). Company B collects the VAT from his customers, but before remitting the VAT to the tax authority company B disappears. Company B is denoted "Missing Trader" (MT).

Figure 2: MTIC fraud - Acquisition fraud



Example

Legal situation	Illegal situation
ICS by Company A: Net price on invoice 100 VAT 0	ICS by Company A: Net price on invoice 100 VAT 0
ICA by Company B: Net price on invoice 100 VAT 20, but deductible	ICA by Company B: Net price on invoice 100 VAT 20, but deductible
DS by honest Company B: Net price on invoice: 95 VAT 19, but the input VAT of 20 is deductible	DS by MT Company B: Net price on invoice 95 VAT 19, but is not paid to tax authorities.
Company's B profit = -5 (loss) i.e. the loss on the sales (95 -/- 100)	Company's B (MT) = 14 (profit) i.e. 19 (VAT not paid) minus 5 (the loss on the sale (95 -/- 100))
VAT = 20 percent	ICA = intra-Community Acquisition ICS = intra-Community Supply DS = Domestic Supply

The VAT loss takes place in MS 2 where the goods are sold by the MT on the national market, usually at a lower price and thus acquiring commercial advantages over other

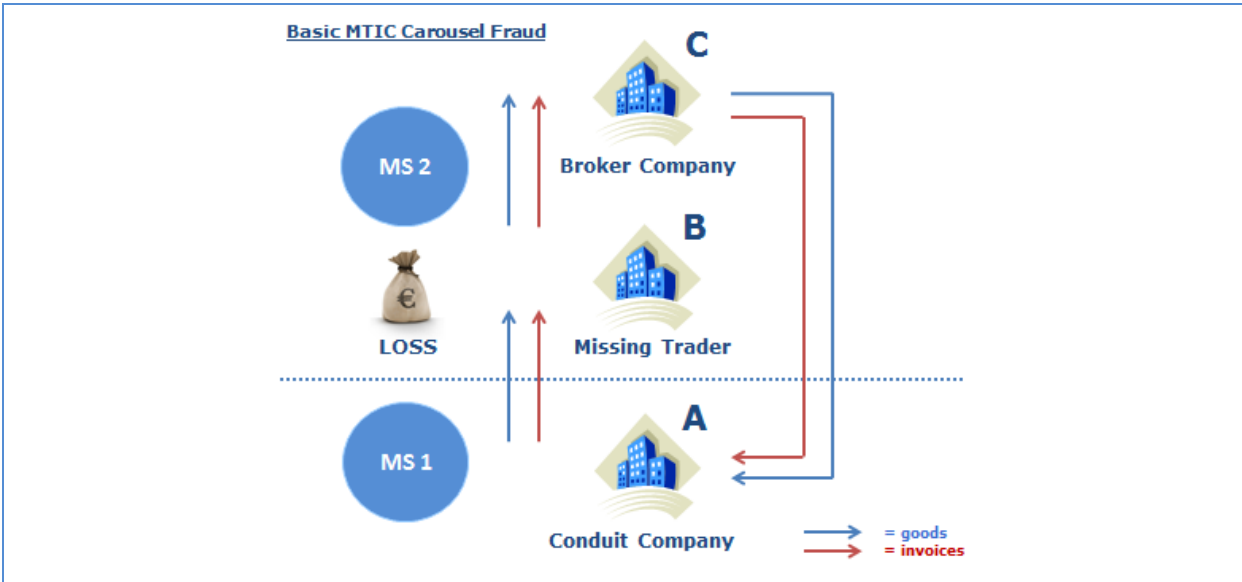
honest traders. If company B is a MT, he will not pay VAT to the authorities and therefore can lower the price and even have more profit. His profit is the charged VAT that is not paid to the tax authority minus the loss on the sale.

2.3.3. Missing trader intra-Community (MTIC) fraud

The scheme of a MTIC fraud, also called "carousel fraud", is an extension of ICA fraud. Instead of supplying goods or services to a final consumer, the supplies are sold back to the original supplier in the other Member State. The scheme is repeated over and over again and therefore looks like a carousel.

In figure 3, Company A (so-called "conduit company") in MS 1 sells goods to company B (MT) in MS 2. The ICS is VAT exempt. The purchase is made without payment of VAT. Company B charges and receives VAT on sales to company C (so-called "broker company"). Company B disappears without having declared and remitted the ICA and the charged VAT to the tax authorities. Subsequently, company C sells the goods back to company A. As the ICS is VAT exempt and he is allowed to deduct the input VAT charged by company B, company C will receive a refund of VAT from the tax authority. When company C makes the ICS to company A the commercial chain is closed and can eventually restart.

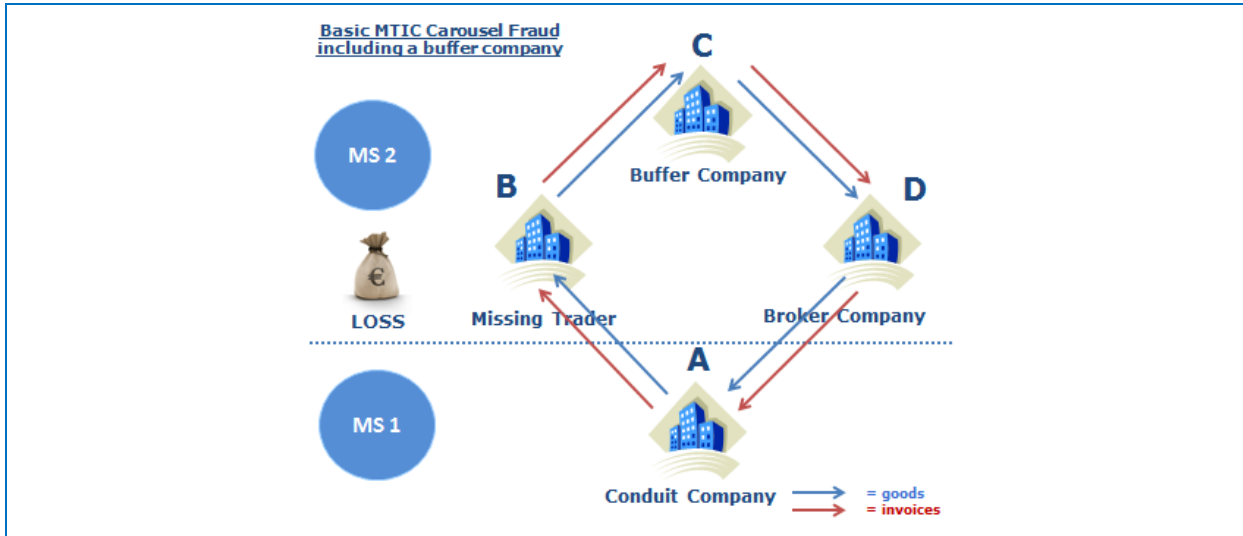
Figure 3: MTIC fraud – Carousel fraud



The scheme may include one or more "buffer" companies to increase its complexity. In that case company B charges and receives VAT on sales to company C (so-called "buffer company"). Company B disappears without having declared and remitted the ICA and the charged VAT to the tax authorities. Subsequently, the buffer company (C) sells the goods to the broker company (company D) charging VAT and remitting balance between input and output VAT to the tax authorities. Subsequently, company D sells the goods back to company A as described in the figure 3. Again, the VAT loss would arise in the Member State of the MT.

The difference between the basic scheme and the scheme in figure 4 is the addition of a buffer company to make the detection of the fraud more difficult.

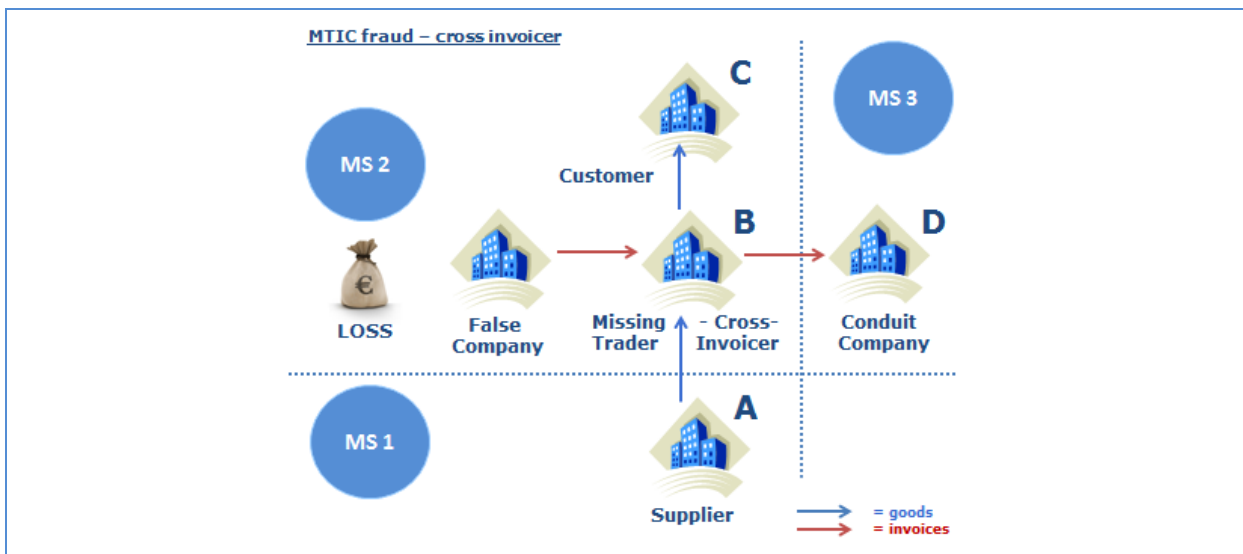
Figure 4: MTIC fraud - Carousel fraud including a buffer company



2.3.4. Missing trader intra-Community (MTIC) fraud – Cross - invoicer

The Cross Invoicer scheme is another form of MTIC fraud. In this case the MT (the cross-invoicer) does not immediately disappear, but rather uses a fictitious invoice, from a false supplier, to deduct the charged VAT, in order to reduce or annul the VAT payments to the tax authority. The domestic supply is typically a fictitious invoice and delivery is made from a fake company or a hijacked company. The onward ICS is also fictitious.

Figure 5: MTIC fraud – Cross invoicer



2.3.5. Missing trader intra community (MTIC) fraud – Contra-trading

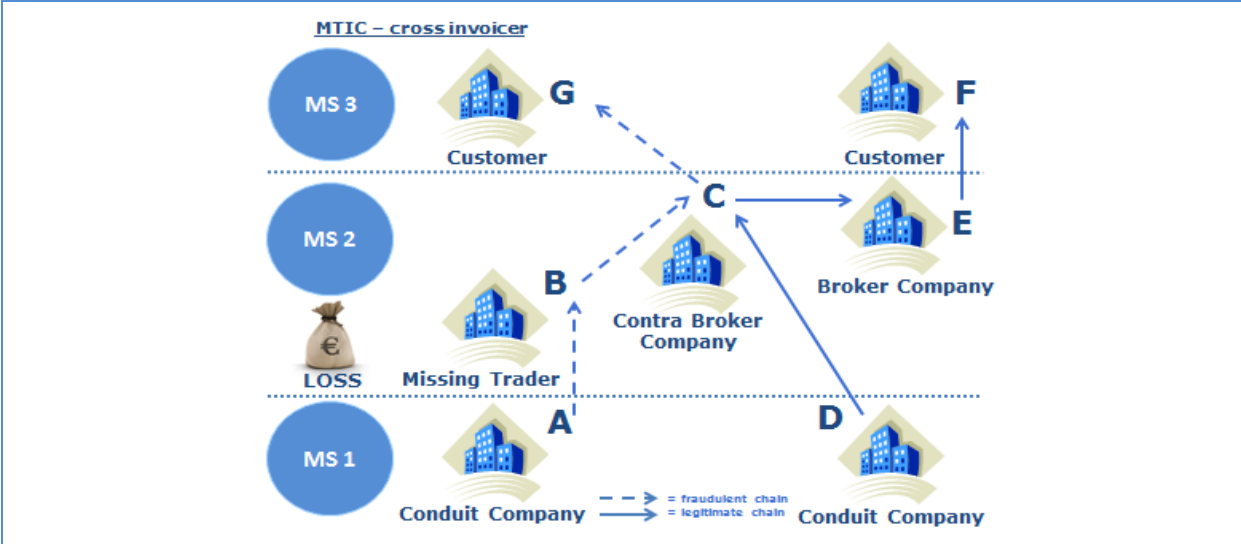
The *Contra-Trading* scheme is a complex scheme in which various transactions chains are introduced to make VAT fraud difficult for the tax authority to discover. The scheme is illustrated in figure 6.

The keystone of the fraud scheme is company C, the contra trader in MS 2. This company makes two acquisitions: (1) a domestic purchase from the MT (company B) which in turn had acquired IC goods from company A in MS 1 without paying VAT; and (2) an ICA from company D in MS1. The MT in MS 2 disappears without paying VAT to the treasury on the goods sold.

Then, the contra trader makes two supplies: (1) an ICS - related to the fraudulent chain - to company G in MS 3; and (2) a domestic supply- not directly related to the fraudulent chain- to company E which in turn makes an ICS to company F in MS 3.

MS 2 suffers VAT losses due to both VAT unpaid by the missing trader and VAT refunds direct, from company C, or indirect, from company E, linked to the fraud.

Figure 6: MTIC fraud – Contra trading

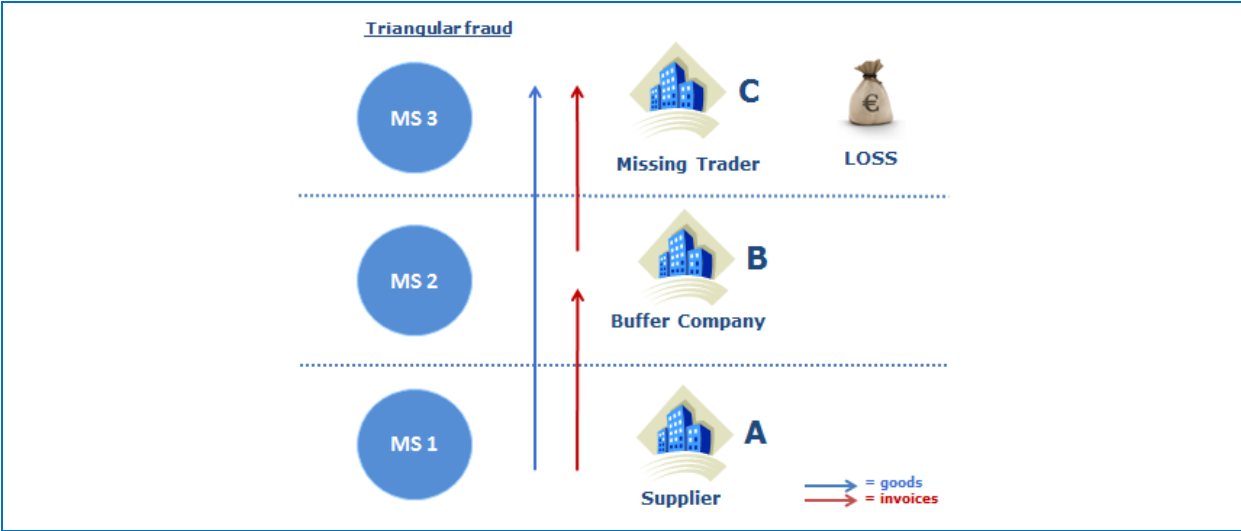


2.3.6. Triangular VAT fraud

In triangular transactions, the goods are delivered directly from the supplier to the final customer. However, the flow of invoices concerns two deliveries. Goods are sold by company A in MS 1 to a buffer company B in MS 2, who sells the goods to the final buyer C in MS 3 and the goods are transported directly from company A to company C. Although there are two VAT deliveries between the companies in the triangular chain, there is only one physical transport of the goods. The fraud occurs by abusing the simplified VAT regime by a MT.

Triangular transactions can take place within one Member State, but also in IC transactions. Triangular transactions are very common in carousel fraud, because the detection of the scheme is difficult.

Figure 7: MTIC fraud – Triangular VAT fraud

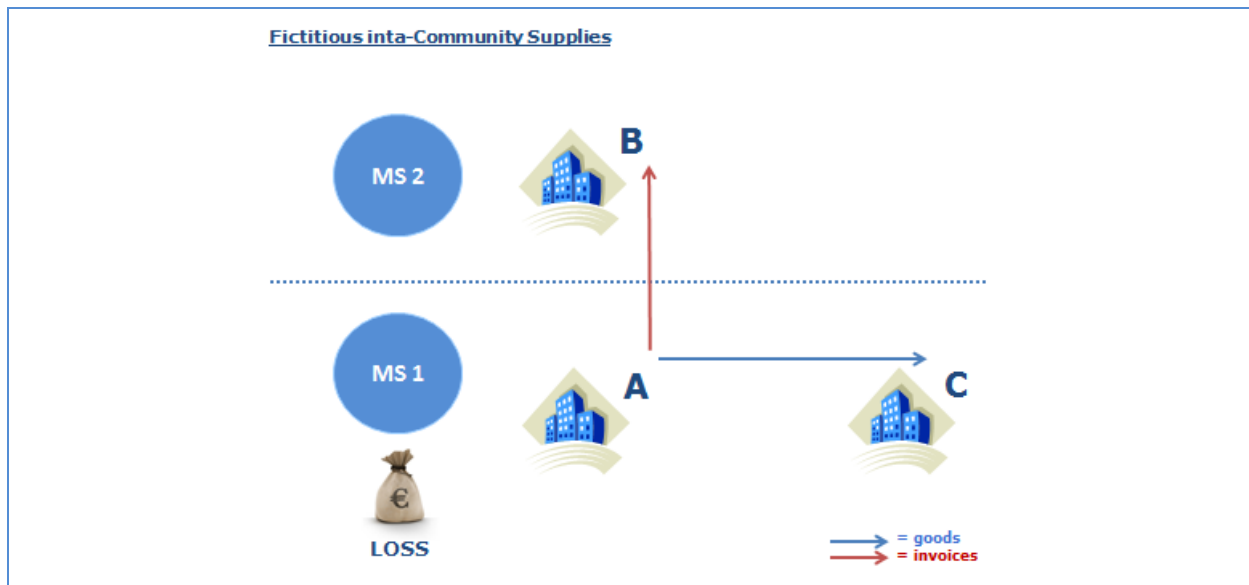


2.3.7. Domestic sales designed as intra-Community supplies

Under this scheme company A in MS 1 simulates an ICS of goods to company B in MS 2. Company B in MS 2 acts like a MT, or the VAT identification number from company B could have been hijacked so that he is not even aware of the fraud.

Company A in MS 1 sells the goods in the domestic market, but does not remit the correspondent VAT as a result of the fictitious IC transaction.

Figure 8: Fictitious intra-Community supplies



3. MTIC in literature

MTIC fraud is a popular topic amongst academics. From different perspectives the issue is addressed, for instance from describing the mechanism of MTIC fraud and proposing solutions, to studies on the VAT gap, trade gap and the VAT system as such. However, it is hard to find research regarding the size and effect of MTIC fraud. This chapter offers a selection of existing literature. It is divided into two parts: 1) general literature about MTIC fraud and 2) literature about estimating VAT fraud, including MTIC fraud.

3.1. General literature about MTIC fraud

Most literature focuses on the VAT fraud schemes, in particular MTIC fraud, and ways how to tackle the problem.

Ainsworth (2010, 2011, 2013) explains in detail the VAT fraud problem (including MTIC) and proposes different tax regulations to prevent fraud. These articles are based mainly on real life fraud cases. In Ainsworth (2013) he summarizes the thirteen most discussed proposals for tackling carousel fraud. A more detailed description of missing trader fraud is presented in Ainsworth (2010). Special cases of tradable services such as voice over internet protocol (VOIP) and CO2 permits are explained. Technology solutions, such as a VAT locator number system and a software certification provision are presented as ways to overcome the MTIC problem. Ainsworth (2011) analyses how to solve the VAT fraud problem by comparing and contrasting three technical proposals of new VAT systems: VLN, RTvat and D-vat.¹²

Similarly, Buskhsh and Weigand (2014) explore similar issues. They describe the different VAT fraud scenarios (acquisition fraud, carousel fraud, contra-trading fraud), followed by a summary of possible solutions and guidelines proposed by various researchers and government agencies. On the one hand, the focus is on procedural changes as a quick reaction to detect the chain of fraudulent transactions (these changes include extended verifications, disruption of criminal activity, cross-border cooperation, reverse charge, scrutiny of new VAT registrations, real-time logging of trades and verification of counter-parties, and collection of VAT in real-time); on the other hand, technical solutions (RTvat, VLN and D-VAT) are highlighted.

Also, Walpole (2014) discusses available options for tackling VAT fraud, such as delaying refunds of VAT, the removal of the zero rate for ICSs, the introduction of the reverse charge mechanism beyond certain specific economic sectors, removing reduced VAT rates or increasing the registration verification. He debates introducing "gold card" schemes, holding parties within a chain of supplies jointly and severally liable for VAT that has not been remitted by other parties in the sequence of supplies, ensuring ultimate recovery of amounts underpaid by means of a guarantee given by a third party or operating a VAT account system. Similarly to others, he considers other options like using D-VAT certification, VAT locator numbers or a lottery system and exchanging and checking the information.

Sergiou (2012) emphasises the nature, extent and cost of MTIC within the EU. The article highlights that, in addition to fully domestic carousel fraud, the EU has seen a rise in cross-border supplies coming from non-EU countries. It underlines that the vulnerability of EU VAT system is due to the fact that the exchange of information between Member States is too slow to expose fraudsters before they have disappeared. Possible solutions should be found in a closer cooperation and understanding among related parties as well

¹² VLN = VAT locator number - business would need to secure a locator number or validate an opposing trader's number when purchasing supplies; RTvat = real-time VAT collection - is about moving the point of taxation from the invoice date to the settlement date; D-vat = digital VAT/certified tax software - introduction and implementation of certified tax software.

as the application of combined restrictive measures, such as extended verification, additional legislation for a more sophisticated VAT system, a generalised reverse charge system to pass the liability for VAT onto the buyer, rather than the seller and ending VAT zero-rating for trade between MS.

Similarly, Van der Hel-Van Dijk and Griffioen (2016) emphasise the importance of administrative cooperation. They present different VAT fraud schemes and analyse the European Commission's measures to tackle them, such as the quick reaction mechanism, the optional application of the VAT reverse charge mechanism, the Eurofisc network and supporting the simplified tax systems. It shows that MS address the problem mainly on a national level. Consequently, VAT-fraud is not resolved at a European level and fraudsters have a free play despite the flood of measures, which causes a VAT gap whose size remains consistently high. The authors stress that a harmonised cooperation between Member States is necessary to be able to combine the available information to unravel the fraud structures and include the examination of the whole supply chain in an EU wide approach. Moreover, a cross-border criminal approach is still possible via a Joint Investigation Team. In addition, root-cause analysis in different economic sectors would be useful to address problems properly and seek for tailor-made solutions.

3.2. Literature about the estimation of VAT fraud

There is hardly any literature about estimating MTIC fraud. Although titles may refer to MTIC fraud, it appears that the content is rather about the VAT gap¹³, shadow economy or VAT fraud in general than MTIC fraud.

One of the first attempts to calculate the size of MTIC fraud is recorded in Ruffles et al. (2003). It focuses on fraud with mobile phones and computer components in the United Kingdom (UK) market. In the UK Ruffles et al. worked closely together with Her Majesty Revenue and Customs (HMRC) who provided them with a valuable insight into MTIC fraud. For confidentiality reasons the UK estimation methodology for MTIC fraud is not included in the article.

Gebauer and Parsche (2003) begin by presenting the traditional method of estimation through the top-down approach of potential VAT revenue, based on national accounts data, input-output tables and special statistics. A problem in estimating VAT evasion by national accounts is the availability of the necessary data. Official statistics are highly aggregated. As a result, details that are relevant when making the estimations are not always evident from the publications. Therefore, some assumptions must be made, which will highly depend on the quality of the available data.

Keen and Smith (2007) propose three ways to produce an estimate of VAT fraud; (1) an estimation based on national accounts (i.e. the VAT gap), (2) a method applied in the UK by HMRC (not available), and (3) other tax authorities' internal approaches.

Fedeli and Forte (2008) look at the gains from VAT evasion for the firms involved in the chain as a consequence of the VAT system in international trade. They explore the rule for sharing the gains from tax evasion between fraudsters and the capture of the market share for which the firms organise themselves into a carousel to exploit the VAT system. The different types of VAT frauds are then formalised into a model and assessed on the effects on the prices of the goods exchanged and on the earnings of the fraudsters participating in the scheme.

Frunza et al. (2010) uses Arbitrage Pricing Theory (APT) models to provide evidence of tax fraud on the carbon allowances market. APT is a theory of asset pricing which assumes that the expected return of a financial asset can be modelled as a linear

¹³ The VAT gap is the difference between the theoretical VAT and the actually collected VAT. The gap includes mistakes, bankruptcies, insolvencies and VAT fraud.

function of macro-economic factors, where sensitivity to changes in each factor is represented by a factor-specific beta coefficient. The authors discovered that the accuracy of the model diminished (was corresponding less to fundamentals) when it was applied for estimating MTIC fraud as prices were not established by markets, but by scheme traders.

Borselli (2011) refers to two main methods for the estimation of VAT fraud: the direct and the indirect approaches. The direct approach is mainly based on data from auditing activities. It mostly applies to statistical inference on non-random data that also need to take properly into account tax risk indicators to avoid misleading estimates. The indirect approach is based on the fact that National Accounts figures are biased since they are based on the VAT Intrastat declarations, and therefore asymmetries in the intra EU trade balance may arise. In this report direct approaches are called bottom-up, while the indirect approach is categorized as a top-down method.

An important summary of research on the VAT gap is done by Barbone et al. (2012). The authors tried to find a relationship between the VAT gap and the VAT burden by using an instrumental variable regression. However, the only statistically strong relationship that can be found is between the size of the VAT gap and the perceived level of corruption in a country. As expected, perceived lower corruption is associated with a lower VAT gap. The authors also point to the bottom-up methodology for estimating VAT fraud. This methodology is confidential due to its use in criminal investigations and therefore not publicly available.

David and Semerád (2014) evaluate existing standard methods used for measuring tax evasion. These methods were not suitable for determining the amount of tax evasion in the fuel sector and therefore they designed a method of quantifying tax evasion on the basis of data concerning distributor prices and information obtained for, the survey of interested entities.

Another view is to take intra-EU trade as a starting point for estimating VAT fraud.

Gradeva (2014) estimates the responsiveness of the trade gap to changes in the VAT rate for the time period of 2004-2009 to see whether there is a significant correlation between changes in the VAT rate and the trade gap. In her work she matches tradeable products to a six-digit level using the available European database on cross-border EU trade. Then the trade gap is defined as the difference between the value of exports to seven new EU MS reported by EU-15 and the value of imports declared in the Eastern European countries of the same product flow.

Frunza (2016) uses the trade gap to estimate MTIC fraud for EU Member States. He employs a panel regression in order to link the amount of VAT collected across EU members to their respective trade gap. This econometric model is then combined with an option pricing model in order to get the MTIC fraud figure. The gap is modelled as a vanilla call option on the imports in a particular country. To our knowledge this is the first econometric model estimating MTIC fraud with macro data. The method is described in more detail in chapter 4.2.2.

In the framework of the EC proposal to modify the VAT system Ernst & Young (2015) studied the impact of the destination principle on MTIC fraud. For the estimation of VAT fraud, a MS survey and a review of available economic and business data was undertaken. An analysis to quantify the size of the VAT fraud gap was made based on the replies of MS Tax Authorities. Built on a limited number of replies (9 MS) on average 20% of the overall VAT gap was considered to be due to MTIC fraud. The estimated weighted average based on the overall VAT gap proportion was 24%. Based on a total VAT gap in 2011 of EUR 193 billion for 26 EU Member States, the VAT fraud gap was calculated to EUR 46 billion for 2011.

Some national tax authorities make estimations of MTIC fraud. For instance HMRC is using a top-down approach to estimate the VAT gap and then a bottom-up methodology

to estimate the magnitude of the VAT evasion. It is not published due to confidentiality reasons.

The Estonian Tax and Customs Board (ETCB) has developed a methodology on how to estimate various types of VAT fraud. The tax authorities are already monitoring the highest risk taxpayers. The ETCB uses bottom-up models for individual compliance risks and assess the level of organized MTIC repayment fraud as by far the largest compliance risk, having caused very significant revenue losses. Also, for reasons of confidentiality the method is not revealed.

The Belgian, Bulgarian and Finnish tax authorities designed an estimation method for MTIC fraud. These are described in chapter 4.

4. MTIC fraud estimation methods

4.1. Introduction

The estimation of MTIC fraud is important for Member States as it helps to evaluate the amount of their VAT revenue losses due to fraudulent behaviour of their taxpayers. However, because of the complexity of the schemes it is almost impossible to identify the size of VAT losses through MTIC fraud and the allocation of those losses in the different Member States. There is a serious risk of double counting in cross border transactions, in particular when there is a difference between the invoice trail of the supply and the trail of the goods.¹⁴

This risk hampers the interpretation of a MTIC fraud figure. Losses identified in one Member State, may have been wrongly allocated. Therefore, it is risky to compare MTIC fraud figures between Member States. Secondly, one cannot just add the figures of the separate EU Member States together to one EU figure, as it will be highly probable that it includes double calculations. It is possible to avoid that kind of double counting only by the use of conventions or common rules through which it is assessed where the fiscal losses would appear.¹⁵

Despite these limitations it is useful to identify the size of MTIC fraud for policy reasons, to show the citizens that MTIC fraud is not acceptable and the government is taking measures to combat MTIC fraud.

In this section, the current available methods, both in Member States and in applied economic research, are briefly summarized. Most methods discussed represent useful approaches to assess the scale of MTIC frauds. Which method to implement among those described is related to the data availability, and estimation tools and techniques already employed. In any case, Member States are encouraged to test and use different methods to compare the results.

The listed methods are grouped in two main approaches, i.e. top-down and bottom-up approaches. They represent, to the best of our knowledge, the state of the art of the methods that might be adopted.

4.2. Top-down approaches

Top-down or macro approaches are estimation methods that allow deducing MTIC fraud estimates mainly from the use of aggregate indicators. Usually aggregated data come from National Accounts or from databases containing aggregated intra-Community transactions or from internal administrative sources (e.g. data from tax declarations).¹⁶ Since these approaches apply macro data, an exact and accurate MTIC fraud estimate is hard to derive. These are worthwhile methods enabling to signal possible anomalies in the economic flows that cause MTIC frauds. The estimation strategy mainly relies on measuring variances of the selected economic variables or on analysing the links between such variables and warning indicators relevant for frauds. The resulting estimates are indicative and when these methods are applied for MTIC frauds estimates, they require a clear indication of the working hypotheses under which the estimates are derived.

¹⁴ For example: when on paper goods are sold from Member State 1 to Member State 2, but in fact the goods are transported directly from Member State 1 to Member State 3. If the buyer in Member State 2 is a missing trader, Member State 2 will consider the loss in its country. However, the real loss is in Member State 3, where the goods arrived and maybe sold on the black market.

¹⁵ For instance, in the bottom-up approach the loss takes place in the MS where the missing trader is established.

¹⁶ A description of possible data sources can be found in chapter 5.

This section will describe two available methods in top-down approaches, i.e. (1) Intrastat mirror statistics and (2) A method based on the variation of macro variables.

4.2.1. Intrastat mirror statistics

In the framework of an assessment whether the introduction of the VAT reverse charge mechanism limited VAT fraud in a certain sector, the Centre for Social and Economic Research (CASE) developed a model using foreign trade data to estimate the size of VAT fraud (CASE 2015). The model is based on measuring the difference between intra-Community Acquisitions (ICA) registered in the MS of the buyer and intra-Community Supplies (ICS) registered in the MS of seller, called mirror statistics. The data on ICAs declared by companies from one MS are compared with the data on ICSs declared by trading partner companies from another MS.

In order to find some indications enabling a specification of concrete types of goods to be investigated further, it is needed at the beginning of the analytical work to study carefully the fluctuation in production, consumption and prices. Significant fluctuation can be partly explained by the changes in economic activities of companies or their surroundings (e.g. economic crisis, special investments etc.). But it may also indicate some irregularities caused by tax fraud (in big scale).

The CASE (2015) analysis distinguishes two types of fraud activities that negatively influence VAT incomes.

- The first type of fraud activities are related to trade operations with the participation of missing traders, who import commodities, usually previously exported or intra-Community supplied, not declaring it for Intrastat¹⁷ and not paying the VAT due. This type MTIC-fraud can be distinguished in a one off MTIC fraud (in the case of a single illegal operation) and repeated MTIC fraud (in the case of multiple use of the same intermediaries, so called brokers and buffers) and carousel fraud (repeated MTIC fraud based on a multiple trade of the same commodity).¹⁸
- The second type is linked to the shadow economy and will not be described, because it is outside the scope of this report.

The above mentioned categorization has been introduced because of the different ways of VAT evasion and VAT fraud, which require different analytical methods to estimate the scale of the fraud.

The CASE (2015) model relies on two assumptions: (1) the missing trader does not register intra-Community acquisitions in Intrastat, since his existence is too short, and (2) the broker has no reason not to register its supply in Intrastat.

The analysis is based on a comparison of discrepancies between ICA reported by the MS of "importation" and ICS reported by the country/ies of "exportation". In order to avert the errors due to burden of the registrations in Intrastat system, the calculation of the differences was not based directly on the Intrastat records' discrepancies, but on their changes over time before and after introducing the reverse charge VAT mechanism.

The observed substantial and persistent differences between registered import in a MS of import and export to the MS of import, reported by EU MS of export, may indicate existence of VAT carousels.

¹⁷ Intrastat is a system that collects statistics on goods traded among Member States. A detailed description of Intrastat can be found in Chapter 5.

¹⁸ For a more detailed description of the fraud types, see chapter 2.

What allows to distinguish between the "repetitive" and the "individual / single" fraud of the missing trader - MTIC is the participation of the same companies in several operations. The estimation of the scale of this fraud is possible because foreign brokers who re-export the product to the MTIC fulfil their obligation to record the transaction in the Intrastat system. This is required for continuation of repeated fraud activity scheme.

However, applying this model can lead to a possible underestimation due to:

- Possible two directions of fraudulent intra-Community trade flows
- Missing Trader could register its acquisition
- Brokers may not register their supply
- MTIC may include extra-Community trade flows with low duties on imports
- The existence of a threshold in the register obligation in Intrastat
- The exemption to declare some kind of goods
- The lack of obligation to record services in Intrastat and
- The existence of different thresholds for each MS

4.2.2. Method based on the variation of macro variables

Frunza (2016)¹⁹ has designed a macro-econometric model using panel regression to estimate the MTIC fraud component of the VAT gap. The model links the VAT collected to the trade gap. The model does not cover VAT fraud linked to the black market and extra-Union traders.

The idea behind the model is when a sudden increase in ICAs is observed, which cannot be explained by economic trends and does not lead to an increase in VAT collected, this may be linked to MTIC fraud. The main assumption in this model is that VAT liability depends on the level of imports. In our case, it means intra-Community acquisitions.

A formula can be determined in order to estimate the VAT gap resulting from MTIC fraud, which is similar to the value of a vanilla call option²⁰. This value can be calculated using a standard formula, where the strike level²¹ is the value of imports at a normal level given an economic framework, and the spot price²² is the value of imports (ICAs) observed (registered).

$$\mathbf{MTIC\ gap = k \cdot Call (Imports, Imports\ normal)}$$

To get an estimation of the level of the MTIC fraud gap, it is needed to know the values of k and the expression Call (Imports, Imports normal). The value of imports at the normal level is estimated through an algorithm (a recursive optimization) using as proxy the reported VAT gap by the EC and assuming that a proportion of the reported VAT gap corresponds to the MTIC. The value of parameter k can be estimated from the equation:

$$\frac{\partial \mathbf{VAT_{observed}}}{\partial \mathbf{Imports}} = \mathbf{\alpha - k \cdot \Delta Call (Imports, Imports\ normal)}$$

¹⁹ Marius-Cristian Frunza, Costs of MTIC VAT fraud for European Union members, April 4, 2016.

²⁰ A vanilla option is a financial instrument that gives the holder the right, but not the obligation, to buy or sell an underlying asset, security or currency at a predetermined price within a given timeframe. A vanilla option is a normal call or put option that has no special or unusual features. The owner of a call has the right, but not the obligation, to buy the underlying instrument at the strike price; the owner of a put has the right, but not the obligation, to sell the instrument at the strike price.

²¹ In finance, the strike price (or exercise price) of an option is the fixed price at which the owner of the option can buy (in the case of a call), or sell (in the case of a put), the underlying security or commodity.

²² A spot price is the current price at which a particular security can be bought or sold at a specified time and place.

The sensitivity of VAT observed (collected by a country) is simply the variation of VAT observed with respect to variation of imports (ICAs), α is the standard rate of the VAT, and delta Δ is similar to the delta of a European call option which can be calculated using a standard formula. If the level of imports is in the range of the trigger level (imports at the normal level given a specific economic framework), the delta could be approximated with the value 0.5.

The model starts with the estimation of the theoretical VAT based on ICAs. The assumption is that VAT theoretical is considered as an increasing function depending on the amount of ICAs. In other words, if ICAs are rising, the VAT collected should increase. It is assumed that this relationship is linear and could be written as:

$$\text{VAT theoretical} \approx \alpha \cdot \text{Imports (ICAs)}$$

Arithmetically, when imports (ICAs) overpass the normal economic level noted with $\text{Imports}_{\text{normal}}$ (the trigger level) the difference between the observed imports and trigger level is the VAT gap resulting from MTIC fraud. Hence, the MTIC gap is proportional to the difference between the actual imports and normal imports:

$$\text{MTIC gap} \approx k \cdot \max(\text{Imports} - \text{Imports}_{\text{normal}})$$

In a fraud free world, the VAT collected by a country would be an increasing function of Imports. The VAT gap increases with the amount of Imports in a country when this amount overpasses the trigger level of Imports. The observed VAT (effective amount collected) increases linearly until the level of Import equals to $\text{Imports}_{\text{normal}}$. When this threshold is overpassed, the observed VAT remains almost constant. The sensitivity of the observed VAT goes towards zero when the value of Imports is much higher than the trigger value (i.e. $\text{Imports}_{\text{normal}}$).

The main driver of the VAT amount from IC transactions perspective is the balance between the Imports and Exports resumed in the Trade gap. VAT collected should be negatively related to the trade gap, as increasing of imports relatively to the exports should induce an increase of VAT collected and vice versa.

The advantage of the model described is the availability of data. All variables could be found using the Eurostat database. The frequency of the dataset is yearly.

According to Frunza the usability of the model is limited due to the fact that it deals with macro data and relies on the assumptions that previous to 2009 MTIC was less.

4.3. Bottom-up approaches

Bottom-up approaches provide MTIC fraud estimates on the bases of the examination of individual data from various sources, such as tax returns, risk registers or third-party information and audit data. These procedures apply various statistical techniques to determine the incidence and value of non-compliance from MTIC frauds. The methods usually consist of two main steps. In the first part, the results of the analysis that identify potential MTIC fraudsters are applied. Secondly, outcomes from the first step require to be extrapolated to determine the extent of non-compliance from MTIC frauds. Extrapolating results to the whole population may follow various approaches. The section below shows methods that are currently applied.

4.3.1. Measuring MTIC fraud based on risk analysis (Belgium)

4.3.1.1. Profiling ICA Cross-Checking and the predictive modelling

The MTIC fraud mechanism necessarily implies at some point undeclared ICAs. Hence, it should be possible to tackle that kind of fraud by observing discrepancies between the ICA figures each MS receives from other MS (L1QD)²³ on the one hand, and the operations taxpayers declare—or fail to declare—in their national VAT returns on the other hand. Many MS provide a statistics box for ICAs on their national VAT return while other MS can infer the volume of ICAs by performing appropriate calculations on the available boxes.

Naturally ICA discrepancies can have many origins aside from deliberate MTIC fraud, such as bad data quality, errors or timing issues. The challenge for tax authorities is therefore to identify and discard ICA discrepancies of a non-fraudulent nature so that they can focus on the fraudulent ones.

Belgium uses a **predictive model** to detect fraud signals in mass data. This model takes advantage of the recent advances in the field of machine learning and the evolution of the computing capacities available today at a reasonable cost. The model predicts a fraud probability for taxpayers based on their features, which mostly consist of past and present L1QD data and VAT returns—as well as various indicators and statistics derived from them. In some cases, adding demographic or geographic characteristics can increase the predictive performance of the model.

Building a predictive model means that the machine learns from past cases, both positive and negative, to predict future cases. If only positive—or only negative—cases are provided, then the machine has no way of learning how to distinguish them. Positive and negative cases should be available in sufficient and not-too unbalanced numbers. The quantity of cases required to attain a decent predictive performance also depends on the quantity of considered features—more features requiring more cases.

Typical models of this kind include decision trees, random forests, logistic regressions and neural networks. Combining models is also possible and often improves the overall performance. However, it is frequently objected that most models lack interpretability, which can raise issues in some circumstances—for instance, when the tax authority is compelled to formally justify the reasons why it decides to investigate a specific case.

The answer to this objection is that it is always possible, for any model, to justify ex-post the reasons why a specific case was chosen by conducting a sensitivity analysis on the taxpayer features. On the other hand, if a global set of business rules is required rather than case-by-case justification, then it is possible to build a decision tree—which has a high interpretability—that will mimic the optimal model prediction and provide the business rules.

It should be stressed that there exists no universal risk score model that would fit all MS at all times. MS do not dispose of the same information and each MS's business context has its own specificities. Therefore, it is the job of the data scientist at national level to build, evaluate and maintain the best possible model to maximize predictive performance within the given geographical, time scope and legal system in force. Once such a model is specified, it should be regularly evaluated and, if necessary, adjusted as time goes by to make sure it remains the best possible one.

²³ L1QD (Level 1 Quarterly Delivery) = aggregated data of ICAs from another Member State

4.3.1.2. The Belgian Experiment

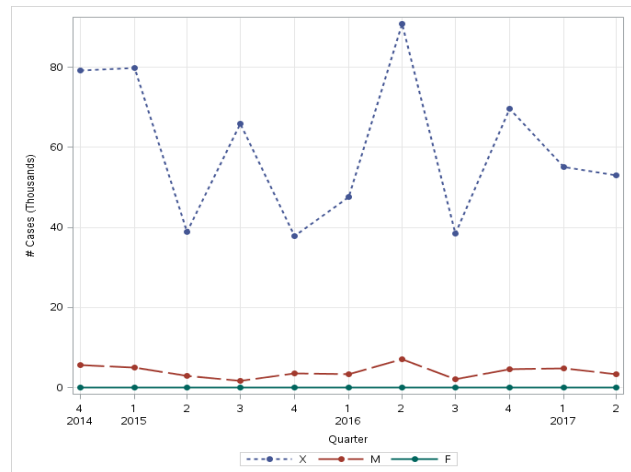
The Belgian Special Tax Inspectorate is currently developing a predictive model to improve MTIC fraud detection among taxpayers displaying positive ICA discrepancies, that is, cases where the ICAs as reported in L1QD exceed the ICAs as reported in VAT returns. Only goods and triangular transactions are considered in the ICA cross-checking.

The preliminary model presented here is a decision tree. By examining past data, the tree highlighted four taxpayer profiles displaying a high probability of fraud. Focussing attention on the selected profiles rather than on the whole set of cases greatly reduces the amount of time and business resources necessary for performing a manual check (see Figure 9).

Figure 9 displays the total number of cases for which the cross-checking of L1QD data and VAT returns detects a positive ICA discrepancy (blue dotted "X" curve), the number of cases corresponding to one of the four profiles the model identifies (red dashed "M" curve) and the number of confirmed MTIC-fraud cases after investigation (green solid "F" curve). The correlation between the "M" and the "F" curves equals 0.34.

Figure 9: Cases with positive ICA discrepancy

Quarter	# Cases (X)	# Cases (M)	# Cases (F)	Model Gain
2014:4	79,080	5,587	64	92.9%
2015:1	79,838	4,989	65	93.8%
2015:2	38,950	2,863	45	92.6%
2015:3	65,836	1,686	50	97.4%
2015:4	37,867	3,559	41	90.6%
2016:1	47,575	3,351	32	93.0%
2016:2	90,818	7,093	66	92.2%
2016:3	38,417	2,020	41	94.7%
2016:4	69,684	4,628	32	93.4%
2017:1	55,160	4,746	10	91.4%
2017:2	53,055	3,276	0	93.8%



In every quarter the model reduces by more than 90% the number of cases to examine based on the ICA cross-checking, which is represented by the vertical difference between "X" and "M". However, even if more than 90% of the ICA cross-checking cases are discarded, there might still be too many remaining cases to be checked manually. Besides, the discarded cases might include fraudulent cases wrongly labelled as non-fraudulent by the model—the so-called "false negatives". These two objections are dealt with as follows.

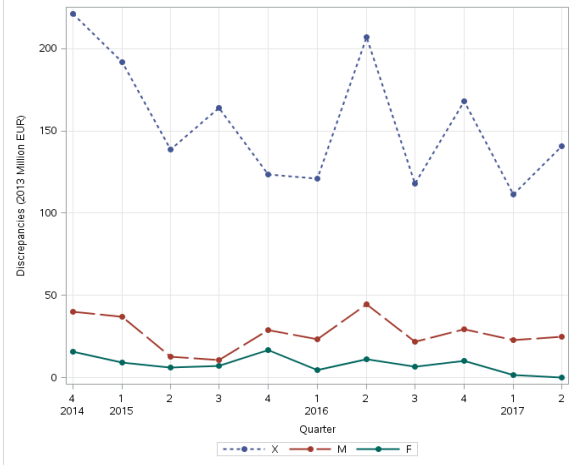
The first step is to reduce as much as possible the number of cases to be checked manually and to make better use of auditing resources. Graphically, this corresponds to lowering the red dashed "M" curve towards the green "F" curve. This can be achieved in two ways:

1. Improving the model: by testing alternative specifications and by working on better and/or more relevant data, the data scientist can increase—up to a point—the predictive power of the model, which would reduce both the number of false positives (non-fraudulent cases identified in the ICA cross-checking) and the number of false negatives (fraudulent cases discarded as non-fraudulent).
2. Increasing the critical score: the model outputs a score representing the MTIC-fraud probability of the considered case, and the business manager is responsible for setting the threshold above which this score is "high enough" for the case to be manually checked. Increasing this critical score would reduce the total number of cases to check as much as desired, but at the cost of increasing the number of false negatives, i.e. fraudulent cases escaping detection. There is a trade-off.

Figure 10 displays the total amount of positive ICA discrepancies for the whole set of taxpayers (blue dotted "X" curve), for the model-identified cases (red dashed "M" curve) and for the confirmed cases of MTIC fraud (green solid "F" curve). The correlation between "M" and "F" equals 0.53 and is statistically significant at the 10%-level.

Figure 10: Total of positive ICA discrepancies

Quarter	Discrepancies (X) - EUR	Discrepancies (M) - EUR	Discrepancies (F) - EUR	Model Gain
2014:4	221,092,083	40,072,787	15,739,954	81.9%
2015:1	191,679,330	36,846,157	8,978,728	80.8%
2015:2	138,456,173	12,497,060	6,379,668	91.0%
2015:3	164,051,150	10,767,216	7,145,160	93.4%
2015:4	123,255,066	28,823,340	16,777,303	76.6%
2016:1	121,063,862	23,264,844	4,746,812	80.8%
2016:2	206,793,771	44,760,511	11,266,908	78.4%
2016:3	117,840,487	22,058,087	6,787,370	81.3%
2016:4	168,105,773	29,586,162	10,059,312	82.4%
2017:1	111,157,506	22,665,371	1,478,044	79.6%
2017:2	140,626,105	24,667,879	0	82.5%



From the profiling part it is possible to derive two main indicators, (1) the predictive indicator of the theoretical amount of the fraud and (2) the effective outcome indicator. The first indicator could be obtained by adding the individual results of the amount of fraud predicted by the model. The second indicator, dependant on the number of audits made, could also be estimated by adding the audit results of the predictive indicator. However, each indicator is estimated using different uplift factors.

4.3.2. Accrued tax by risky people (Bulgaria)

4.3.2.1. Methodology – selecting the risky persons –risk register

The method is based on the assumption that the amount of VAT fraud in the VAT gap is actually the VAT which was accrued or has to be accrued from supplies by persons registered in the "risky taxpayer register" to non-risky persons (persons who do not appear in this list). This approach should be seen as a **general approximation to VAT fraud where MTIC fraud is included**. For this estimation method transaction level data is needed about sales and purchases, which is not available in every country.

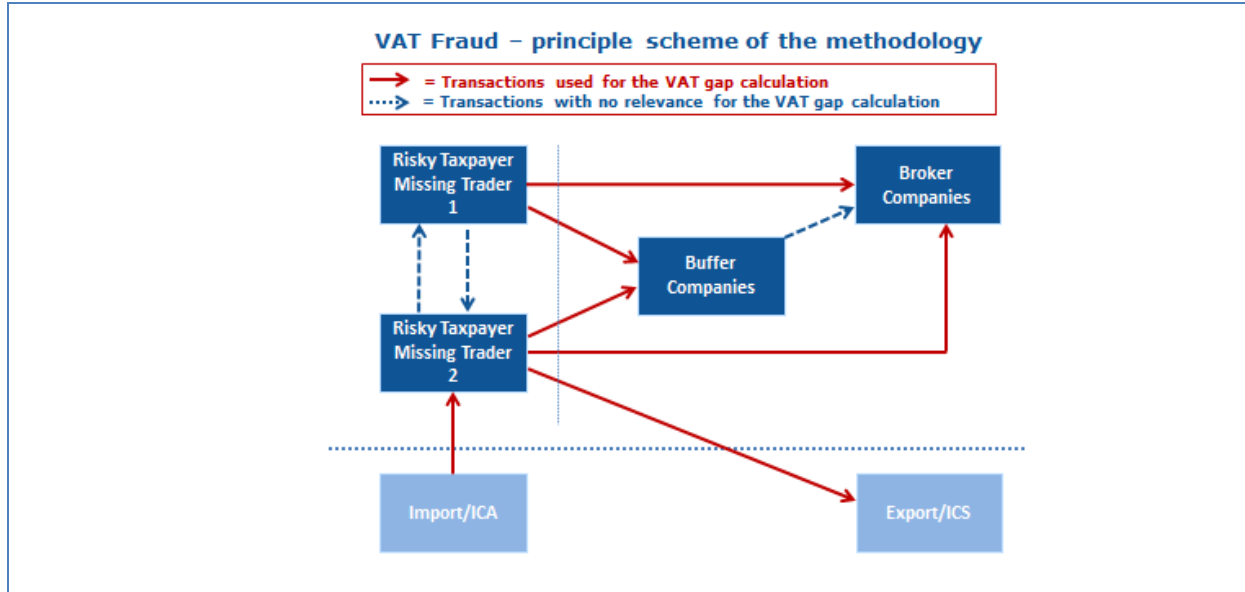
The first step is to establish and update the list of risky persons about whom the tax administration has serious suspicions that they are involved in organized VAT fraud. For this purpose the following actions are taken:

- All persons and companies related to the persons from the current list have been identified and checked and if they match the risk criteria added to the register of risky persons and companies;
- The counterparties of all persons from the list, different from the suppliers already existing in the list, have been identified. A review of those persons (with regards to the purchases, sales, non-paid liabilities, certificates of audits, etc.) has been made and after defining them as risky persons involved in VAT frauds they have been included in the register of risky persons;
- All persons deregistered under the VAT Act on the initiative of the tax authority have been compared to the complete new list of risky persons. The persons who have not been included in the list were reviewed in detail again and added to the register if applicable.
- Hence there is a continuous updating of the list of risky persons.

4.3.2.2. Measurement of the tax gap from VAT frauds: Sales method

For the calculation of the VAT fraud gap all transactions between risky taxpayers who appear in the risk register (missing traders) and non-risky taxpayers are taken into account (see the red arrows in figure 11).

Figure 11: Principle scheme of the methodology



For all these transactions the sum of the tax base and accrued VAT on the supplies is calculated for all persons in the register and for each of the periods in the following cross-sections:

- According to the type of transaction with regard to its tax treatment (taxable, exempt, intra-community acquisition, intra-community delivery, etc.).
- These sums, differentiated according to the type of delivery, are divided into:
 - Type of the document (invoice, protocol, credit and debit note, customs declaration).
- The so received amounts according to the type of document are divided into:
 - "recorded" and "unrecorded" by the deliveries' receivers (the recorded and unrecorded status of the transactions is a result from the automated matching in order to find congruence between the declared transactions' sales and purchases by the persons registered under the VAT act).

The tax gap from VAT frauds has been calculated, before making any corrections and after assessing the sums according to the abovementioned cross-sections, using the following formula:

$$\text{VAT FGAP} = \text{VAT}_T + \text{VATR} * \text{TB}_0 - \text{VAT}_{\text{KNm}} - \text{VAT}_{\text{Import}} + \text{K}$$

VAT FGAP	=	VAT fraud gap
VAT_T	=	VAT accrued from all taxable supplies, without the tax base of the 0% rate supplies;
VATR	=	VAT rate in %;
TB₀	=	Tax base of the supplies taxable with 0% rate /export, ICS, etc. pursuant to the VAT Act;
VAT_{KNm}	=	VAT with regard to issued credit notes reported by the counterpart;
VAT_{Import}	=	VAT accrued from import;
K	=	Correction for VAT gap from undiscovered risky persons;

The amount of **K** is the part of the tax gap caused by probably existing but undiscovered risky persons. The correction amount is calculated on the basis of the number of undiscovered risky persons and the average amount of the VAT-gap caused by the ones already identified and included in the risky persons register.

The number of undiscovered risky persons is determined on the basis of newly identified risky persons after reviewing a random sample of all other VAT-registered persons.

As long as the methodology is based on individual data of the persons and companies involved in VAT frauds and data for specific transactions, it enables not only the calculation of the gap, but also the acquisition of knowledge about various aspects of fraud as a whole.

For example the, in this way, calculated VAT fraud gap is divided according to various cross sections as:

- The type of tax payers involved in VAT fraud – sole traders, limited liable companies, joint-stock companies, etc.;
- The fraud subject – which goods or services are most often used in VAT fraud;
- The type of affected transactions – domestic, ICA, ICS, triangular transactions, import, export, etc.;
- Territorial distribution of the VAT-frauds – distribution of the risky persons, the VAT-gap and the companies that benefit from VAT frauds according to the tax administration's offices;
- Risky person's "life duration" – distribution and tendencies of the risky person, according the numbers of tax-periods they have been active.

This new knowledge enables the tax administration to build on an effective anti-fraud strategy and to distribute and direct its resources to counteract this phenomenon.

4.3.2.3. Sources of information

The calculations are based on data from the available databases in the tax administration over the period 2005-2017, collected and generalized at national level. The sources include:

- Filed annual tax returns under the Corporate Income Tax Act /CITA/ and the Personal Income Tax Act /PITA containing identification data about the persons, the type of operations, etc.;
- Filed VAT returns;
- Monthly filed purchases and sales registers in electronic format under the VAT Act by every VAT liable person – containing structured data about the transactions the person has completed: according to tax period, type of the document (invoice, protocol, credit and debit note, customs declaration), type of the transaction with regard to its tax treatment (taxable, exempt, intra-community acquisition, intra-community delivery, etc.), identification data about the counterpart, tax base and occurred tax for the delivery, etc.;
- Results from automated matching in order to find congruence between the declared transactions' sales and purchases;
- The reports of the results from audits where liabilities under the VAT Act for the abovementioned period were audited – containing the amount of the additionally established liabilities, etc.;
- The register of risky persons and companies, supported by the National Revenue Agency since 2000, where there is information proving their participation in the commitment of tax frauds – containing: identification data about the person and companies, the type of risky person /missing trader, buffer, broker/, the type of fraud, the subject of fraud, etc.

4.3.3. Measuring MTIC fraud based on risk analysis (Finland)

This section presents a summary of the calculation of the MTIC fraud gap in Finland in the years from 2011 to 2014.²⁴ The MTIC gap is calculated by a bottom-up method, which includes three steps:

1. Profiling and finding fraudulent companies
2. Calculating the number of fraudulent companies
3. Estimating the gap

The MTIC scheme is complex and includes many different forms of fraud and roles of fraudsters. In this study the focus is on intra-Community acquisition fraud (see chapter 2.32. for an explanation of the scheme). Other schemes were not represented well enough among the discovered fraudulent companies considering the needs of statistical inference. Even the number of discovered Intra-Community Acquirers is dangerously low, and causes problems when estimating the number of fraudulent companies.

The estimated number of discovered fraudulent Intra-Community Acquirers (F-ICA) is about 100 per year and the annual MTIC gap from their actions is about 12–15 million euros. The size of total annual MTIC fraud varies between 12 and 35 million euros.

4.3.3.1. Selection or Profiling method

The corner stone of profiling is a group of so called discovered fraudulent companies found in tax audits. There are two essential groups of enterprises

- Discovered fraudulent intra-Community acquirers , missing and defaulting traders found in tax audits (36 observations in 2011–2014)
- The population, other companies that have intra-Community acquisitions (230 000 observations in 2011–2014)²⁵

In profiling, the features of discovered F-ICA will be compared with the same features in the population. Companies, whose features are similar to those of discovered F-ICA, will be selected from the population.

The idea of profiling is to define features that are more frequently found in the group of discovered F-ICA than in the population. Therefore, a feature will make a difference between fraudulent and other companies. For a company, a feature is true if its value exceeds or, in some cases, falls below, the limit set for the feature. We call this a hit.

It is normal that F-ICAs have some features in common, but at the same time some features are different. Complex definitions are necessary to exclude false hits, but the problem is that complex definitions will exclude some fraudulent companies. The profiling method allows for different combinations of features: large numbers of features are included in the model, yet the combination of features can vary. The number of features used in profiling was 30.²⁶

The two essential concepts of profiling are the probability and the score.

- The probability of a feature in a group is the number of hits divided by the total number:

$$P_{\text{feat}_p} = N_{\text{feat}_p} / N_p \quad \text{and} \quad P_{\text{feat}_d} = N_{\text{feat}_d} / N_d , \text{ where}$$

N_{feat_d} and N_{feat_p} = the number of hits; N_d = the number of D-F-ICA; and N_p = the size of population

²⁴ This paper is a summary of a report by The Grey Economy Information Unit / Finnish Tax Administration.

²⁵ A low limit of 1.000 euros per year; sources of information are tax and VIES reports.

²⁶ For example (not limited): VIES > VAT (goods), increasing VIES, no payments to Tax Administration, no turnover, no wages, no fixed assets, short life span company, responsible persons change.

- The score of a feature includes the probabilities and the numbers of hits in both groups:

$$P_{\text{feat_d}} \times (1 - P_{\text{feat_p}}) \times (N_{\text{feat_d}} / N_{\text{feat_p}}), \text{ where}$$

- $P_{\text{feat_d}}$ = probability, D-F-ICA → impact on the score ↑
- $P_{\text{feat_p}}$ = probability, population → impact on the score ↓
- $N_{\text{feat_d}}$ = number of hits, D-F-ICA → impact on the score ↑
- $N_{\text{feat_p}}$ = number of hits, population → impact on the score ↓

The **company score** is the sum of scores of those features where a company has a hit.

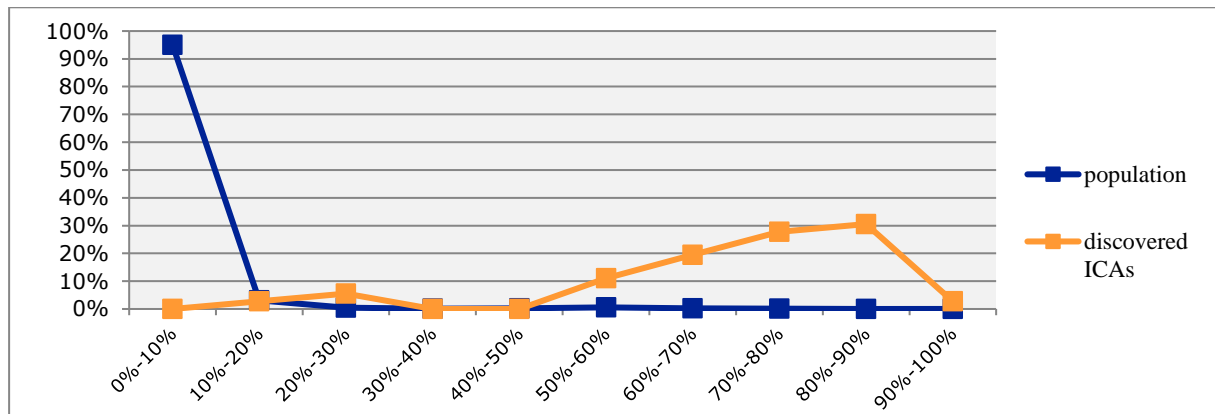
The **relative company score** is the company score divided by the sum of scores of all features.

We call the sum of scores of all features the **strength of model**.

There are deciles of relative company scores that are used to calculate distributions. A **distribution** is the number of companies in each relative class: 0%–10%, 10%–20%, ..., 90%–100%.

The shapes of the distributions shown in figure 12 are very different from one another: 98 percent of population fall in the first two groups of distribution (between 0% and 20%), while 97 percent of discovered F-ICA fall in the eight remaining groups (above 20%).

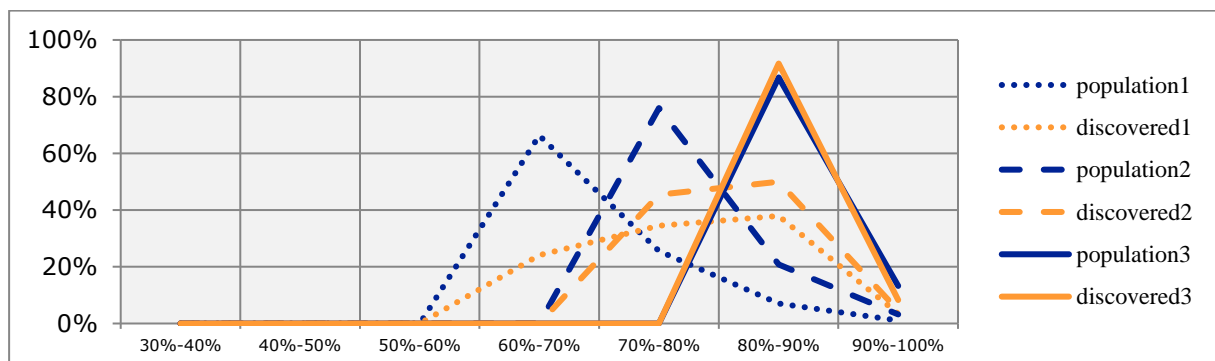
Figure 12: The result of profiling



4.3.3.2. Estimation method

It is interesting to find a sub-group of the population that has a distribution similar to D-F-ICA. This group would be similar to discovered F-ICA with respect to the features in the model. Therefore, companies in such a sub-group would be F-ICAs.

Figure 13: Cut distributions



If bottom classes are cut out, distributions become more like one another. There are three different cuts shown in figure 13. It is assumed that the reason why shapes become more similar as classes are cut out, is that the proportion of fraudulent companies in the population becomes higher.

The ideal cut-off point is the one that separates a group of the population that includes only F-ICAs from the rest of the population. This point may be found by comparing the distributions of the two groups. It is found where the shapes of distributions become similar enough (some element of chance should be allowed).

There are alternative ways of comparing distributions. The success of comparing distributions depends on how well the profiling was done. The stronger the model, the easier it is to find a good cutting point. Also the stronger the model the less the cut-off point influences the estimated number of F-ICAs in the population.

When cutting out bottom classes, also some F-ICAs (in both groups) are cut out. Yet, they must be included in the final estimated number. For example, if the distributions are cut at 80%, 12 discovered intra-Community acquirers and 68 companies of the population will be included in the cut. The share of the cut of all discovered intra-Community acquirers is 33 percent. It is assumed that the same proportion of F-ICAs was included in the cut of the population. Therefore, the number of F-ICAs in the population is estimated to be 204 ($\approx 68 / 0,33$).

Table 1: The estimation of the number of F-ICAs in the population (X)

	F-ICA	Population
Included in the cut 80%–100%	12 (choose)	68 (choose)
Included in the cut 80%–100% / all	33 % (divide 12/36)	33 % (assume)
The number of F-ICAs	36 (known)	X = 204 (estimated)

The number of F-ICAs was calculated in four different ways: (1) the total and using stratifications, (2) size, (3) VAT type and (4) the combination of size and VAT type. The size was measured by goods purchases registered in VIES. There were three different VAT types: missing, incomplete or short, and correct VAT declarations.

The final estimated number was reached by comparing different results and companies above the cut out limits in different stratifications. The final estimated number of fraudulent ICAs is 339–389 companies in 2011–2014. Together with discovered F-ICAs, this makes approximately 100 F-ICAs per year.

Table 2: The estimated number of fraudulent intra-Community acquirers

	F-ICA	Population	F-ICA	Population
The number of fraudulent intra-Community acquirers in the population	339–389	345–360	356–378	339–389

After estimating the number of fraudulent companies, the calculation of the tax gap is straight forward: the amount of VIES-purchases of the selected companies is multiplied by the VAT rate.

The way of calculating the tax gap includes four assumptions:

- VIES reports are correct;
- Fraudulent companies do not pay taxes;
- All reported purchases will arrive in Finland;
- The whole amount of fraud has been detected in discovered F-ICAs.

VAT reports are likely to be less reliable than VIES reports. The seller does not always know about the fraud, and even if he knows, he has an incentive to report the sales (the deductions).

It was discovered from the data that the payments of discovered F-ICAs to the Finnish Tax Administration were minimal, in fact quite often they equalled to zero. This fact supports the second assumption.

The last assumption means that the possibility of remote missing traders is ignored: the goods reported in one country, will in fact go to another country. The tax loss happens in the country where the goods arrive in. The simple reason why remote missing traders are ignored, is that the lack of data to measure this phenomenon.

At this moment losses caused by other actions, for example intra-Community supplies, cannot be estimated. However, the theoretical maximum can be calculated by assuming that all F-ICAs were part of a carousel fraud.

4.3.4. Tax gap extrapolation from random audit data

One bottom-up method of estimating the tax gap is using random audit data. In this section the method is adapted to the MTIC case. At its simplest, this method requires that the tax administration takes the findings (e.g. underreporting, mistakes) from the audits it is performing on a randomly selected population, and extrapolates these to the whole population to get the tax gap.

4.3.4.1. Data collection

Pure random audits as an estimation method works best when there are many fraudulent companies whose fraudulent activities are similar in scale to each other. Thus, the method is not useful in the case of the biggest tax payers.

Before starting with random sampling it is important to discuss the aim of the sample. Beside different tax gap estimations there are other possible goals too, such as having the threat of an audit on every taxpayer. A key question for tax gap estimation is whether the level of the tax gap or its evolution in time should be estimated.

To have a reliable estimate of the tax gap, minimum requirements of data collection should be met. These are:

- Periodically collected, well-structured data of a certain quality;
- The availability of detailed declarations' data;
- Detailed audit or desk check reports to be available in a database;
- Detailed data for compliant and non-compliant companies too.

It is especially important to have all relevant and available data sources connected in the database by taxpayer. Available data sources are very different by tax administration and include third party information as well. All data is relevant that might be related to the chance of companies being fraudulent or to the possible scale of their fraudulent activities.

Depending on the area to be estimated and the available resources and timeframe, another important decision is the method of random sampling and the size of the sample. It can have significant costs to conduct enough random audits to have estimates with a good precision. Without going into detail on sampling techniques it is important that every tax administration should select the version that best fits its preferences.

4.3.4.2. Extrapolation

The main advantage of random audits is that they do not include any distortion on their sample. This makes extrapolation easy. As taxpayers selected for random audits are similar to not selected ones, it can be assumed that a similar amount of evaded tax would have been found at the non-audited companies. The tax gap can be calculated by a simple multiplication. In the case of a more complicated sampling technique it is not that simple, but the essence of the method is the same.

Extrapolation is a much more challenging task in the case of non-random sampling. The selection process should be modelled with statistical models or by using information about the actual selection process or by a combination of the two. This model should be regularly tested and developed with special attention to changes in the selection process. With the model extrapolation to the whole population can be achieved. It is useful to have multiple alternative methods to have a robust picture about the evolution of the tax gap.

4.3.4.3. Selection random samples for MTIC estimation and stratification

MTIC fraud is usually conducted by a small number of companies who evade huge amounts of VAT. This means that a general random audit is not a good method for estimating the scale of MTIC fraud. Selected random audits on a subgroup of high risk companies (for instance companies involved in IC trading with no employees and/ or created recently or other specific features) is a more precise and cost effective way of estimation. Whether the selection of such a subgroup is possible, depends on the available data and it needs careful evaluation. This section describes main issues which should be considered when putting together an estimation method for the MTIC fraud gap.

While for random audit and other selection methods usually a company is treated as the unit of analysis, MTIC fraud is conducted by a group of companies. If connections between companies are not used in the selection and extrapolation processes, the model can be inaccurate.

When calculating tax gap from audit data it is important to measure the real tax loss. In organised MTIC fraud the same tax loss can occur at multiple companies of a carousel network and adding them together would be an overestimation.

It is very important to measure the accuracy of the selection process. At the same time it is dangerous to over fit the model to a specific measure of accuracy as it can lead to a methodology which selects well a subgroup of fraudulent companies, but is unable to detect other fraudulent companies which are not so similar to them.

4.3.4.4. Advantages and disadvantages

Advantages

Random audit methods enable the tax administration to have a detailed view of the tax gap arising from MTIC broken down by different characteristics: type of behaviour, trade sector, taxpayer type, size, any characteristic that has been used in tax gap estimation or has been used to characterise actual taxpayers, who caused the loss. It enables the tax administration to answer the question who and what causes the MTIC gap or the part of the MTIC gap which was possible to estimate through a bottom-up method.

It may give more detailed information in the way a fraudster behaves in setting up a type of fraud. This can help to redefine the sample used to detect fraud.

Disadvantages

Random audit methods use detailed operational information (e.g. from tax returns), which could mean that unregistered and/or concealed activities ("black" economy) are outside the scope. Depending on the nature of tax evasion, bottom-up methods may not give a comprehensive picture of the MTIC gap.

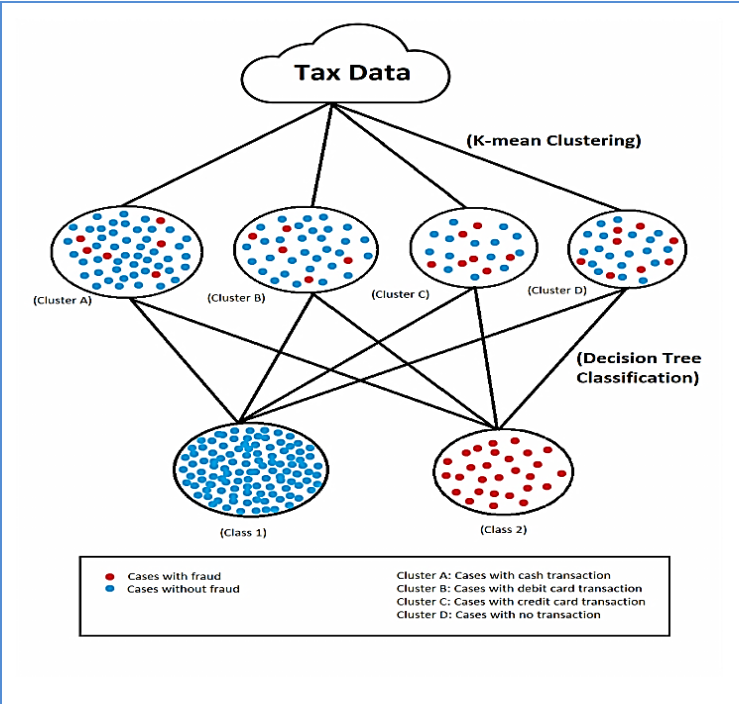
Often random audit methods are based on compliance or enforcement activities which can, in some cases, take a long time to be completed. This means that there can be a delay between the relevant tax year and the availability of the audit results. Particularly complex cases, which are often those involving large tax losses, can take a number of years to resolve, and their data will be finalized only at the end of the process in the sample. The resulting gap estimates typically refer to earlier periods than those from the top-down methods.

As bottom-up methods require detailed information, extra effort is needed to ensure data quality. Whereas the top-down methods rely on external data, the bottom-up methods require tax administrations to take care of the quality of their own data. They need to focus on collecting the right data in the right way. If there is a need to collect data that is only required for tax gap estimation purposes, it requires good cooperation and communication between different units. It will require additional human resources when compared to methods that can be carried out by analysts only. In tax administrations the collection of different types of data is usually organized by different structural units. If the necessary detailed information is not readily available within the tax administration extra effort will be needed to collect this data.

Tax returns, auditing procedures and data sources available for tax administrations are different per country, therefore audit-based tax gap estimations cannot be used for cross-country comparison. A possible approach that has not been tried yet would be to use very limited data sources that are available in every EU country.

4.3.5. K-means clustering

Figure 14: K-means clustering



An option that might be worth exploring in more detail is the method of k-means clustering. The application of this method for anomaly detection is mainly linked to the banking sector and identifying credit card fraud, insurance fraud, and other anomalies in client behaviour. Some countries have used the k-means clustering method to predict fraudulent claims of taxpayers (Wodajo, 2017), or to detect tax evasion (Dias 2015 and Shukla 2018). Recently the method is applied in a Commission study on excise duty (still to be published).

Source: IFERP, vol. 5, Issue 3, March 2018, page 58

Available tax data is at the core of the analysis. After pre-processing the data and feature selection the k-means clustering algorithm is applied to form clusters of taxpayers with similar behaviour. This technique is a kind of unsupervised learning, which is used to find groups in data, where k represents the number of groups (clusters). The k-means clustering algorithm results in centroids of the k-clusters. Each centroid of a cluster is a collection of parameter values which defines the resulting group.

Then decision trees are used to classify each cluster into taxpayers with and without fraud and patterns in their associated behaviour are detected. Using these characteristics and patterns, the artificial neural network is trained and potential fraudsters could be detected based on the information available.

Applying k-means clustering to fraud detection was also briefly discussed in Chandola et al. (2009), and in more detail in Lindholm (2014) and Baesens et al. (2015).

4.4. Comparison of the different methods with regards to possible utilities

Table 3 provides an overview in which the various methods are assessed against six applications, which are:

- The estimation of the MTIC gap at EU level
- The estimation of the MTIC gap at country level
- The comparison of the results between countries
- The comparison of the results year by year
- The breakdown the MTIC gap
- The comparison of the MTIC gap to the total VAT gap

Table 3: Comparison of the different models

Method	Intrastat mirror statistics (CASE)	Method based on the variation of macro variables (Frunza)	Measuring MTIC fraud based on risk analysis (Belgium)	Accrued tax by risky people (Bulgaria)	Measuring MTIC fraud based on risk analysis (Finland)	Random audit	Transactions Network Analysis ²⁷	K-means clustering
EU level MTIC gap estimation	Yes	Yes	No	No	No	No	Yes	No
Country level gap estimation	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country comparison	No	Yes	No	No	No	No	Yes	No
Year by year comparison	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Breakdown of MTIC gap	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Comparable to total VAT gap	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

²⁷ See chapter 6 for a description of the Transaction Network Analysis (TNA)

5. DATA sources

For the estimation of the size of VAT fraud, it is important to identify the relevant data sources. Many sources are available, but not everything is obtainable at EU level. Furthermore and equally important is the quality and reliability of the data. One should take into account that some data is collected for other purposes, which may make the data less useful for estimating the size of MTIC fraud.

Depending on the estimation method, different sources of information can be used. A summary of the main data sources is provided in this chapter.

5.1. National VAT declarations

VAT is declared by taxpayers using VAT returns in the country where they are registered. VAT returns are a necessary source of information for all estimations of VAT fraud. Micro level VAT return data is available only for the National Tax Administrations as a general rule. Aggregated yearly tax income is available on Eurostat. The regulation of VAT returns is very different among Member States. For example there are different rules about the required frequency of declaration (yearly, quarterly or monthly).

5.2. Recapitulative statements and VIES data

Under the current VAT system intra-Community supplies of goods are exempt from VAT in the Member State of dispatch when they are made to a taxable person in another Member State who will account for the VAT on arrival. Therefore any taxable person making such supplies must be able to check quickly and easily that their customers in another Member State are taxable persons and do hold a valid VAT identification number. For that purpose, inter alia, each tax administration maintains an electronic database containing the VAT registration data of its traders. Such information includes the VAT identification number, the trader's name and the trader's address.

On the other hand, any taxable person making intra-Community supplies must submit periodically a so-called recapitulative statement to the local tax administration showing a list of the purchaser's VAT numbers and the trade net turnover value per period.

A computerized VAT Information Exchange System (VIES) was set up to allow for the flow of the data held across the internal frontiers which: (1) enables companies to obtain rapidly confirmation of the VAT numbers of their trading partners and (2) enables VAT administrations to monitor and control the flow of intra-Community trade to detect all kinds of irregularities, including MTIC fraud.

5.3. Intrastat data

Trade operators from EU Member States who sell and transport goods to other clients in other EU Member States are obliged to fill in the Intrastat form. The information included in this mandatory statement feeds INTRASTAT, a system of collecting statistics on goods traded among Member States of the European Union (EU).

Intrastat reports almost all transactions leading to movements of goods in the EU, in particular, (1) exchange of goods entirely acquired in the EU; (2) exchange of goods from countries outside the EU that have been put into free circulation in EU; and (3) exchange of goods from the combinations of (1) and (2). A list of exclusions from reporting goods is also contemplated and concern for instance, emergency aid for disaster areas, some goods for and following temporary use and monetary gold.

Each country involved in the trade makes its own declaration: the suppliers submit the "dispatch Intrastat forms" when the goods are physically dispatched; the purchasers submit the "arrival Intrastat forms" when the goods physically enter the country.

There are some limitations in reporting the EU transactions: one depends on the yearly amount of intra-community goods traded at company level and the other depends on the decision of each Member State about the thresholds of arrivals and dispatches to be declared.

For the purpose of the VAT fraud analysis, the existence of thresholds makes a comparison between EU countries difficult. The Member State that acquires goods may declare a different amount than the Member State that sells them. Therefore, any use of Intrastat data to detect anomalies in the trades among Member States require to account for these limitations. Despite this shortcoming EU countries may use Intrastat data to make a cross check of VAT figures from VAT returns in order to claim insufficiently reported IC transactions.

Intrastat data provides also useful information on goods traded, in particular the value of goods in terms of both the invoice value and the statistical value. Contrary to the taxable amount, the statistical value takes into account also the transport and insurance costs incurred in delivering the goods from the place of their departure to the border of the importing Member States.

Other information concerns, for instance, the unique commodity codes; the Member State of consignment or the Member State of destination on dispatch; the country where the goods are produced or manufactured.

5.4. Audit findings

An important source of information is the result of an audit carried out by the tax administration. The main outcomes presented in the audit report cover tax related facts and corrections in the submitted VAT returns, as well as a number of topics, such as the type of business, the organisation of the company, its owners and managers, relationships with other businesses, number of staff, location and facilities, accountant, bookkeeping system, financial situation and legal issues.

5.5. Other information

Beside data listed above, for a good bottom-up approach other micro data are valuable for estimation purposes as well. Almost any information collected about the taxpayer could be a possible indicator of VAT fraud. Especially important are other tax returns (for instance (corporate) income tax) and information about other (possibly fraudulent) companies of the owners and managers of the company. Also information from Eurofisc²⁸ and third parties (for instance Customs, police, labour inspection, banks and credit card companies) may help in identifying fraudulent behaviour. Some Member States use additional sources of information on VAT liability of taxpayers, such as invoicing software connected to the tax administration or detailed VAT returns, where every partner of the company is listed. Also, information received from the exchange of information between Member States' competent authorities (SCAC requests) may provide for the detection of fraudsters.

From a macro point of view information supplied by the national accounts and budgetary reports might also be useful but it has to be taken into account that this data is made available on a cash basis.

²⁸ Eurofisc is a network for the quick exchange of targeted information on possible fraudulent companies and transactions.

5.6. Overview of data used by the various methods

Table 4 provides an overview of the different data sources that are used by the various top-down and bottom-up estimation methods.

Table 4: Overview of data used by the various methods

Method	Intrastat mirror statistics (CASE)	Method based on the variation of the macro variables	Measuring MTIC fraud based on risk analysis	Accrued tax by risky people	Measuring MTIC fraud based on risk analysis	Random audit	Transactions Network Analysis ²⁹	K-means clustering
National VAT declarations			x	x	x	x		x
Data from recapitulative Statements and VIES			x	x	x	x	x	x
Intrastat data	x							
Audit data			x		x	x		x
Other information (including Eurofisc data)			x	x		x	x	x
National VAT declarations		x						

²⁹ See chapter 6 for a description of the Transaction Network Analysis (TNA)

6. A European approach

Currently, only a few Member States calculate the MTIC fraud gap in their country. For reasons explained in the introduction of chapter 4, adding the available figures together to get an EU figure is not an option. Therefore, the project group discussed the potential of the Transaction Network Analysis (TNA) approach. Despite its bottom-up character, TNA has the characteristics of a European approach for its specificity and in particular the access to disaggregated national data. Although, in principle TNA is a risk score model, it may help in getting more insights in the size of the MTIC gap. Another method that may have potential for estimating MTIC fraud is K-clustering. This method is used for instance in calculating the size of excise fraud.

6.1. The TNA Project

On the basis of the methods described above, assessments can be measured in MS making use of available data. These estimations can then be compared to each other. The TNA project offers an opportunity for a common approach because the available data will be similar for all MS and treated in a structured environment. The analysis can also take place in real time.

6.1.1. Context

The purpose of the project, i.e. the development of a Transaction Network Analysis platform, is to support Eurofisc with a new IT system. As such the scope of the project is to deliver TNA as a fully deployed application ready for further configuration, optimization and testing by Eurofisc.

In this context, the project's scope covers application design, development, testing and deployment. However, the configuration of the system with business specific data analysis algorithms and business rules remains outside the scope of the TNA project. These activities along with the efficient use of the system are expected to be performed by MS in the context of Eurofisc. These external activities are of high importance and necessary to reach the aforementioned expected benefits.

The TNA platform does not seek to go beyond the current Eurofisc framework, e.g. limitations of exchanging targeted information on MTIC fraud. The business opportunity that TNA represents does not target to change the "what", but improve the "how" by addressing the following points:

- Automating the collection of targeted information over the CCN30 network, e.g. VIES;
- Being able to visualize suspicious networks without manual intervention;
- Improving MS ability to share data and qualify traders identified by Eurofisc by making them less dependent upon manual uploads of information by each MS;
- Improving the quality of information shared by enhancing IT facilities for collecting and sharing the information, e.g. by enforcing quality controls;
- Improving the detection rate and accuracy by introducing advanced data analytics, e.g. network analysis algorithms.

³⁰ CCN network means the common platform based on the Common Communication Network and Common System Interface developed by the Commission to ensure all transmissions by electronic means between competent authorities in the area of Customs and Taxation.

6.1.2. Evaluation of MTIC Gap

Based on the data in the TNA, it will be possible to evaluate the losses caused by missing traders. This can be done by summing the ICAs of the defaulters. The data are collected automatically by direct access to the data used for the VIES. Failures are provided by the Eurofisc Liaison Officers (ELO) (feedback from Eurofisc and all MS must be introduced in the TNA) and also on the basis of the analyses.

These statistics can be drawn up on an ongoing basis. These will be available in several sizes, including:

- By injured country
- By sector of activity
- By typology
- By suppliers
- By supplier country

This will also allow the injured country to develop the most powerful indicators based on positive cases. The process must be fully automated.

It should be noted that this approach can only be useful if MS decide in advance on the conditions that have to be fulfilled, such as the quality of data, skilled staff, clear terminology, etc.

7. Conclusions and recommendations

7.1. Conclusions

Estimating the size of MTIC fraud is difficult. The complexity of this type of VAT fraud makes it hard to correctly allocate the revenue losses, which may lead to double counting and thus overestimating the size of the problem.

Despite this risk, it is still worthwhile to calculate the MTIC gap. It will provide an indication of the dimension of the problem and can guide policy makers in designing appropriate measures to tackle MTIC fraud.

Literature and experiences in Member States on how to calculate the MTIC gap are scarce. The project group has identified a few examples, which, as far as they are not revealed for reasons of confidentiality, are described in this report.

From an EU perspective, the analysis of the methods showed that a top-down approach seems to be the most adequate estimation method for the MTIC gap because of the use of macro data. The two approaches (CASE/Frunza) offer opportunities, but more research is needed to make these approaches better applicable for EU purposes.

The *bottom-up approach* has the advantage of applying the information that countries bring from audits and risk analysis, that is, all the relevant information is considered when estimating the gap. The methodologies listed in this report offer the possibility to come up with some figures to have a more precise estimation of the real dimension of MTIC fraud and also allow for the possibility to implement predictive modelling.

The *bottom-up approach* allows for the identification of the different risks causing the revenue loss. It is very important to correctly establish the population to be studied. The main disadvantages are the need for well skilled resources, as it involves a number of audits, processing the results, etc. which makes this approach very costly. The fact that individual country-estimates are not easily adaptable to the EU level is also a disadvantage. Therefore, at this point it is difficult to have an EU level estimation based on a bottom-up approach.

For EU purposes the TNA approach may offer an opportunity to get more insights in the size of the MTIC gap. It has the advantage of preventing double counting of losses and will identify in an automatic way suspicious actions based on established internal rules. However, at this stage the TNA is not operational yet.

7.2. Recommendations

1. National estimations of MTIC fraud either by top-down or bottom-up approaches may provide insight in the size of the MTIC problem and could help policy makers in designing effective anti-fraud measures. Therefore, Member States are invited to estimate the size of MTIC fraud at national level for which the methods mentioned in this report can be used or may be used as a source of inspiration.
2. There is no perfect method to identify the size of MTIC fraud. Member States are trying to find a proper method, but the complexity of the fraud schemes and available data sources make it hard to design a satisfactory method. Therefore, Member States are encouraged to share their experiences and knowledge on how to estimate MTIC fraud to help each other in tackling this phenomenon.
3. The TNA approach may offer a common approach. Member States are encouraged to further invest in this approach by developing specific KPI's and indicators to allow an evaluation at national and EU-level.

4. This report is a first starting point for identifying estimation methods for MTIC fraud. There is still little experience and therefore the project group recommends continuing research and projects in this field at national and EU level.
5. Cooperation with other EU institutions (ESTAT) or external parties (example: universities, statistical offices, law enforcement bodies) could be beneficial.
6. Apart from MTIC fraud, other types of VAT fraud should be estimated to get the full picture of revenue losses based on fraud and evasion. For these fraud schemes other methods might be applied. Therefore, the project group advises to continue the research on VAT fraud estimations.

References

- Ainsworth, R. T. (2010), VAT fraud: MTIC & MTEC-THE tradeable service problem. Working Paper No. 10-39, Boston University School of Law, Boston, USA
- Ainsworth, R. T. (2011), Technology Can Solve MTIC Fraud – VLN, RTvat, D-vat Certification. International VAT Monitor, May/June 2011, p.153-160
- Ainsworth, R. T. (2013), Tackling VAT Fraud: Thirteen Ways Forward. Working Paper No. 13-36, Boston University School of Law, Boston, USA
- Aizenman, J., and Y. Jinjirak (2005), The Collection Efficiency of the VAT: Theory and International Evidence. NBER Working Paper No. 11539
- Barbone, L., M.M. Bird, and J. Vazquez-Caro, (2012), Cost of VAT: A review of the Literature. Case Network Reports no. 106/2012
- Borselli, F. (2008), Pragmatic Policies to Tackle VAT Fraud in the EU. International VAT Monitor, September/October 2008, p.333-342
- Borselli, F. (2011), Organized VAT Fraud: Features, Magnitude, Policy Perspectives. Banca D'Italia, Occasional Paper 106
- Borselli, F., S. Fedeli, and L. Giuriato, (2015), Digital VAT Carousel Fraud: A New Boundary for Criminality. Università di Roma. Working Paper n.170
- Bukhsh, F.A. and H. Weigand, (2014), VAT Fraud: Possible Technical and Ontological Solutions. Tilburg University, 2014
- CASE c.s. (2017), Study and Reports on the VAT Gap in the EU-28 Member States: 2017 Final Report, TAXUD/2015/CC/131, Warsaw, 6 September 2017
- CASE - Centrum Analiz Społeczna Ekonomicznych, (2015), Ocena czy poprzez wprowadzenie od dnia 1 października 2013 r. odwróconego obciążenia VAT na niektóre wyroby stalowe ograniczono nadużycia w rozliczaniu podatku VAT w obrocie tymi towarami oraz czy warunki funkcjonowania przedsiębiorstw w branży objętej zmianami uległy poprawie (Assessment whether the introduction of VAT reverse charge mechanism on 1 October 2013 for certain steel products limited VAT fraud with respect to the trade in these goods and whether the business development environment in the industry under changes has improved), prepared at the request of the Ministry of Finance from Poland.
- COM(2017) 566 final, Communication on an Action Plan on VAT, towards a single EU VAT area - Time to decide, COM(2016) 148 final, Brussels, 7.4.2016 and Communication on the follow-up to the Action Plan on VAT - Towards a single EU VAT area - Time to act, Brussels, 4.10.2017
- COM(2017) 569 final, Proposal for a Council Directive amending Directive 2006/112/EC as regards harmonizing and simplifying certain rules in the value added tax system and introducing the definitive system for the taxation of trade between Member States, Brussels, 4.10.2017
- Dias A., C. Pinto, J. Batista, M.E. Neves. (2016), Signaling tax evasion, financial ratios and cluster analysis, Observatório de Economia e Gestão de Fraude (OBEGEF), Working papers no. 51/2016
- David P. and P. Semerád (2014), Possibilities of measuring tax evasion related to fuel sale, Procedia Economics and Finance 12 (2014) 121 – 129
- Ernst & Young (2015), Implementing the 'destination principle' to intra-EU B2B supplies of goods, Feasibility and economic evaluation study, Final Report, 30 June 2015

European Court of Auditors (2015), Tackling intra-Community VAT fraud: More action needed, no. 24 Special Report.

Fedeli S. and F. Forte (2008), Models of Cross-border VAT Fraud, Società Italiana di Economia Pubblica, Pavia, September 2008

Fedeli S. and F. Forte (2012), A Game Theoretic Approach to Cross-border VAT Evasion within EU Member States and Its Relationship with the Black Economy. Economic Analysis & Policy, Vol. 42 No. 2, September 2012

Finnish Tax Administration (2017), Measuring MTIC Fraud based on risk analysis.

Fiscalis Tax Gap Project Group (2016), THE CONCEPT OF TAX GAPS, Report on VAT Gap Estimations. (FPG/041)

https://ec.europa.eu/taxation_customs/sites/taxation/files/docs/body/tgpg_report_en.pdf

Fiscalis Project Group (2013), Good practice Guide to tackling Intra-Community VAT fraud, updated version 2013.

Frunza M., D. Guegan, and A. Lassoudiere (2010), Statistical evidence of tax fraud on the carbon allowances market, March 11, 2010

Frunza, M. (2016), Cost of the MTIC VAT Fraud for European Union Members, Schwarzthal Kapital, Paris, April 2016.

Gebauer, A. and R. Parsche (2003), Evasion of Value Added Taxes in Europe, IFO approach to estimating the evasion of Value Added Tax on the basis of national accounts data (NAD), CESifo DICE Report 2/2003

Gradeva, K. (2014), VAT Fraud in Intra-EU Trade, Goethe University Frankfurt.

IMF Country Report No. 14/133 – Republic of Estonia - Revenue Administration Gap Analysis Program - The Value-Added Tax Gap (2014)

<https://www.imf.org/external/pubs/ft/scr/2014/cr14133.pdf>

Keen, M. and S. Smith (2007), VAT Fraud and Evasion: What Do We Know, and What Can be Done? IMF Working Paper WP/07/31

Martin, P. and R. Walters (2013), Fraud Risk and the Visibility of Carbon, Journal of Crime and Justice, vol 2, no.2 (2013)

Matthew, K. and J. Lloyd-Williams (2000), Have VAT Rates Reached Their Limit? An Empirical Note. Applied Economics Letters 7, p.111-115

Reckon LLP (2009), Study to Quantify and Analyse the VAT Gap in the EU-25 Member States. Discussion paper

Ruffles, D., G. Tily, D. Caplan and S. Tudor (2003), VAT Missing Trader Intra-Community Fraud - The Effect on Balance of Payments Statistics and UK National Accounts. Economic Trends no. 597

Shukla Y., N. Sidhu, A. Jain, S.T. Sawant-Patil (2018), Big data analytics based approach to tax evasion detection, International Journal of Engineering Research In Computer Science and Engineering (IJERCSE), Volume 5, Issue 3, March 2018

Sergiou, L. (2012), VAT Carousel fraud in the European Union, Journal of Accounting and Management, vol. 2, no. 2, 2012

Waerzeggers, Ch. (2010), VAT in a Common Market: Lessons for the GCC. <http://ssrn.com/abstract=2338139>

Sokolovska, O. (2016), Cross-border VAT frauds and measures to tackle them, MPRA Paper No. 70504

Thackray, M., E. Hutton, and K. Kapoor (2015), Revenue Administration Gap Analysis Program – the Value-Added Tax Gap, Finland, International Monetary Fund, Fiscal Affairs Department U.K., Government, HM Revenue and Customs. Measuring Tax Gaps—2016 (October 2016). <https://www.gov.uk/government/statistics/measuring-tax-gaps>

Van der Hel-van Dijk, L. and M. Griffioen (2016), Tackling VAT-Fraud in Europe: A Complicated International Puzzle, Intertax, vol.44, issues 4, 6 and 7

Walpole M. (2014), Tackling VAT Fraud, International VAT Monitor, September/October 2014. P.268-263

Wodajo B., S. Appalabatia, Telkapalli Murali Krishna (2017), Predicting the fraudulent Claims of Tax Payers; a case of Boditi Town Revenue Authority, SNNPR, Ethiopia, International Journal of Scientific & engineering Research, Volume 8, Issue 10, October 2017

Annex I. - List of participants

Country	Name of participant	Administration
BELGIË / Belgique (Belgium)	Mr Yannic Hulot	Ministry of Finance, Belgium
ČESKÁ REPUBLIKA (Czech Republic)	Mr Jiří Vlach	General Financial Directorate of the Czech Republic
ESPAÑA (Spain)	Ms Estrella Gonzales	Agencia Estatal de Administracion Tributaria
	Mr Ramiro Ledo	
ITALIA (Italy)	Ms Elena D' Agosto	Italian Revenue Agency
LIETUVA (Lithuania)	Ms Agnė Jakubauskate	State Tax Inspectorate under the Ministry of Finance of the Republic of Lithuania
MAGYARORSZÁG (Hungary)	Mr Ván Bálint	Ministry for National Economy, Hungary
NEDERLAND (Netherlands)	Mr Menno Griffioen	Tax and Customs Administration, Netherlands
	Mr Lou Verberkt	
POLSKA (Poland)	Mr Tomasz Mazur	Ministry of Finance, Poland
	Ms Teresa Chomiczewska	
PORTUGAL (Portugal)	Ms. Carla Pereira Rodrigues	Centre for Tax and Customs Studies Portuguese Tax Authority
ROMÂNIA (Romania)	Mr Victor Ogneru	National Tax Administration, Romania
SLOVENSKÁ Rep. (Slovakia)	Ms Martina Motkova	Ministry of Finance of Slovak Republic
SUOMI (Finland)	Mr Aki Savolainen	Finnish Tax Administration
EUROPEAN COMMISSION	Ms Jannetje Bussink	Directorate General Taxation and Customs Union
	Mr Tanel Puetsep	

Annex II. - List of presentations

First meeting: 29 September 2016, Brussels (BE)

- Ms Jannetje Bussink (DG TAXUD): Context, background and aim of the VAT subgroup
- Ms Jannetje Bussink (DG TAXUD): VAT fraud – Missing Trader Intra Community Fraud
- Mr Marius Voicu (DG TAXUD): VIES system (VAT Information Exchange System)
- Dr Marius-Cristian Frunza (guest speaker): Cost of MTIC VAT fraud for EU Members
- Mr Yannic Hulot (member of VAT subgroup, Belgium): Estimating MTIC fraud – Belgium approach

Second meeting: 13 January 2017, Amsterdam (NL)

- Ms Iona Stretton, Mr Kishen Patel (guest speakers, Ernst & Young): Estimating the size of VAT fraud

Third meeting: 3 May 2017, Utrecht (NL)

- Ms Kirsi Ristola (guest speaker, Finland): VAT fraud: Profiling and tax gap estimation – Finish approach

Fourth meeting: 13 July 2017, Utrecht (NL)

- No presentations

Fifth meeting: 10-11 October 2017, Lisbon (PT)

- Mr Nikolay Petkov (guest speaker, National Revenue Agency of Bulgaria): GAP from VAT Fraud – How to Understand? - Bulgarian experience
- Mr Adam Fajkusz (guest speaker, National Tax and Customs Administration Hungary): Applying statistical-mathematical modelling in tackling VAT fraud. Tax difference model.
- Ms Catrina Boogh-Dahlberg (guest speaker, EC, DG ESTAT): National Accounts perspective - Exhaustiveness and VAT Fraud

Sixth meeting: 12 December 2017, Brussels (BE)

- No presentations

Seventh meeting: 30-31 January 2018, Lisbon (PT)

- No presentations

