Identifying Business Misreporting in VAT Using Network Analysis

Christian González-Martel[a], Juan M. Hernández[a,b,*], Casiano A. Manrique-de-Lara-Peñate[c]

[a] Department of Quantitative Methods in Economics and Management, ULPGC
[b] University Institute of Tourism and Sustainable Economic Development (TIDES), ULPGC
[c] Department of Applied Economic Analysis, TIDES, ULPGC

This document is the Accepted Manuscript version of a Published Work that appeared in final form in *Decision Support Systems*.

To access the final edited and published work see
https://doi.org/10.1016/j.dss.2020.113464

Highlights

Identifying Business Misreporting in VAT Using Network Analysis

Christian González-Martel,Juan M. Hernández,Casiano A. Manrique-de-Lara-Peñate

- A model to determine factors influencing errors in VAT declarations is presented.
- Four types of errors when declaring taxes are analyzed.
- The role of relational factors in declaration mismatches is highlighted.
- Firms with VAT declaration errors in a specific direction are usually interrelated.

Identifying Business Misreporting in VAT Using Network Analysis

Christian González-Martel[a], Juan M. Hernández[a,b] and Casiano A. Manrique-de-Lara-Peñate[c]

[a]*Department of Quantitative Methods in Economics and Management, University of Las Palmas de Gran Canaria, c/Saulo Torón, 4, 35017 Las Palmas, Spain*

[b]*University Institute of Tourism and Sustainable Economic Development (TIDES), University of Las Palmas de Gran Canaria, c/Saulo Torón, 4, 35017 Las Palmas, Spain*

[c]*Department of Applied Economic Analysis, University Institute of Tourism and Sustainable Economic Development, University of Las Palmas de Gran Canaria, c/Saulo Torón, 4, 35017 Las Palmas, Spain*

ARTICLE INFO

Keywords:
Networks
fraud detection
VAT declaration
random forest

ABSTRACT

Efficient detection of incorrectly filed tax returns is one of the main tasks of tax agencies. Value added tax (VAT) legislation requires buyers and sellers to communicate any exchanges that exceed a certain amount. Both statements should coincide, but sometimes the seller/buyer and its counterpart declare different amounts. This paper presents a method to detect those businesses that are more prone to misreport in their VAT declaration. Using the information of such declarations for a region in Spain during year 2002, we generated a transaction network formed by the tax declarations of buyers and sellers. Four types of errors were assigned to each business in the network, defined from the mismatch between the amount declared by the firm in question and its partners. We applied a random forest algorithm to detect which firm-based and which network-based characteristics influence each error type. The results show the importance of relational factors among businesses in determining the probability of presenting VAT declaration errors. This information can be used to promote more efficient inspections.

1. Introduction

A value-added tax (VAT) is a tax on consumption that first rose to prominence in the 1960s and is now one of the main sources of indirect taxes in most countries in the world. The most common procedure for the implementation of VAT consists in sellers charging the tax to their clients. If the client is itself another company, it can get a refund of the VAT paid to suppliers against the VAT they have to charge and subsequently pay to the revenue service for their own sales. At the end of the tax period, the final tax return submitted by each agent broadly comprises the difference

ORCID(s):

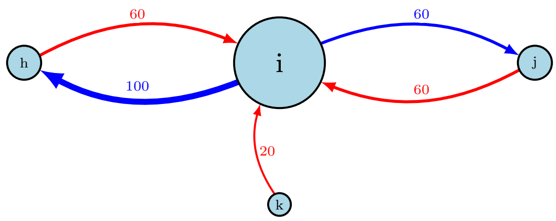

Figure 1: Example of VAT declaration network.

between the total taxes passed on to its buyers and the total taxes paid to its suppliers. This explains why most tax administrations oblige firms to declare their purchases and sales, at least those whose operations exceed a certain threshold over the course of a year, with an explicit identification of the counterparts.

As with other taxes, VAT fraud can be related to many behaviors, including failure to register, misclassification of commodities, taxes collected but not remitted, collusion between agents, and the creation of fictitious activities for the sole aim of asking for refunds [16]. However, the way in which the tax functions, as described above, makes declaring less sales, higher acquisitions, or both, the most specific sources of VAT fraud. In fact, although considered a success in terms of its collection capacity, its refunding system opens the door to all types of fraud both in domestic and international trade transactions [17].

To illustrate how the system works, Figure 1 shows a simple example of a VAT declaration network. Red arrows indicate sales declarations and blue arrows purchase declarations. A single business can operate as both buyer and seller (i.e. i, j and h). As shown in the figure, both statements agree in the relationship between i and j, but not in the other relationships: the declared amount of i and h disagree, and while i does not declare any purchases from k the latter firm declares sales to i. The aim of this paper is to identify the features in a VAT declaration network that best characterize tax statement mismatches. Some of the studied features correspond to the characteristics of the firm in question (economic branch, size, etc.), but so-called relational factors also play a role in the characterization of mismatches. For example, contagious effects (firms mimic partner's behavior) and homophily (firms are prone to be connected with similar firms) are expected to influence business behavior regarding VAT declarations. In this regard, the complex network perspective [4] provides useful metrics to represent these and other features in large networks such as those formed by VAT declarations in a regional economy. Additionally, data mining techniques are suitable to deal with empirical network data [27].

One of the major challenges when studying VAT declaration mismatches is that in general it is not known which of the tax filers has made an incorrect declaration. Therefore, we cannot directly assign a fraud label to either of the two firms by only looking at their VAT declarations. To deal with this issue, instead of trying to identify fraudulent firms, we define four different types of VAT declaration errors that a firm can commit according to its role in the transaction (seller or buyer) and the sign of the mismatch. These errors depend on the counterpart's declaration but do not predetermine any fraudulent intention by either of the two agents. Then, we apply a data mining algorithm (random forest) to detect the features of the firm and the features of the network that influence the propensity to each type of specific error in the VAT declaration. More specifically, the paper answers the following questions: is there any characteristic that is typical of incorrect VAT tax filers? Is there any self-organized structure of these incorrect VAT tax filing firms? For example, is there any kind of homophily among them? By answering these questions, we provide information about the characteristics that influence misreporting, leaving to the VAT experts the questions of interpretation and appropriate action.

The model was applied to a VAT declaration network in a region of Spain (Canary Islands) in the year 2002. Among other findings, the results show a contagious effect in the transaction network, in the sense that firms are more prone to file incorrect VAT returns when they are related with other incorrect tax filers than with correct tax filers.

Summarizing, the major contributions of the paper are:

- We define four types of statement errors in VAT declarations.
- We find network-based features that explain errors in VAT declarations.

- We provide a model to estimate firm and network features that characterize each type of error. This model can be used by tax agencies to detect sectors or specific groups of tax filers which are more prone to commit mistakes in VAT declarations.

The rest of the paper is organized as follows. The following section presents previous studies related to the topic of this paper. Section 3 presents the data source, exploratory results and the methodology for data analysis. Section 4 presents the random forest (RF) results and the last section reports the conclusions of the study.

2. Related work

This paper is related to previous studies dealing with financial fraud detection which have used data mining. Financial fraud includes a series of criminal behaviors such as money laundering, credit card and social security fraud, etc. In recent decades, multiple data mining techniques have been applied to fraud detection [22, 26]. Among them, classification algorithms are commonly used. These methods try to identify features that describe suspicious firms, and start from a pre-defined model. The models are trained with real financial data with which some firms and their transactions have been previously detected as fraudulent. This problem faces several challenges given that fraud is uncommon and generally hidden among correct behavior. Several specific data mining techniques belong to this group, including logistic regression, support vector machines and RF. Of these, RF has commonly obtained good performance in fraud detection [2].

Traditionally, the models used for fraud detection take into account exclusively local (intrinsic) factors, which refer to characteristics of the firm [25, 9], leaving aside relational effects among firms that may also influence fraud. However, several experiments analyzing the explanatory factors of tax fraud have shown the existence of reciprocity between agents. In particular, [5] identified the simultaneous existence of different levels of reciprocity, with "strong reciprocity" the most common. Under this concept, agents tend to behave more correctly if they are treated correctly and more incorrectly if they are treated incorrectly. Applying this behavior to taxpayers, the tendency to evade tax obligations will rise or fall depending on whether the behavior that they observe by the other agents is more or less disadvantageous to them and iniquitous. This result is contrary to the predictions of a self-interest model [14], where agents follow in the footsteps of potential future material gains. [11] use survey data to check the validity of the conditional cooperation hypothesis, finding strong empirical support for the influence of other taxpayers in the tax morale and evasion decision of individual taxpayers.

We can find another perspective of horizontal factors in the way tax privacy affects tax evasion. [3] designed an experiment in which they tested the level of tax evasion under three levels of tax privacy: no information on tax behavior, anonymous information of individual tax behavior and complete information of tax evaders. Two opposite attitudes can be expected after public disclosure of taxpayers' behavior, contagion and shame. The first of these effects initially appears in the second level of tax privacy, while the shame effect only appears in the third one. The authors confirm the existence of both effects, but the shame effect appears to dominate only in the first periods of the experiment whereas the contagion effect dominates in the latter periods of the experiment and overrides the shame effect. In other words, knowing what others do strongly influences agents' attitudes towards tax evasion.

[8] present another experiment in order to incorporate two other non-economic explanations of tax compliance, empathy and sympathy. They consider that empathy makes you share others' feelings while sympathy is considered an emotional concern about others' wellbeing that does not necessarily coincide with their emotions. Although their results are not entirely conclusive, they support the idea that non-economic elements also explain tax fraud behavior, and that moral issues play a significant role in explaining behavior.

In this respect, recent financial fraud detection models have included the role of relational factors in the models, obtaining more accurate results than when using only local variables [e.g. 20, 13, 24, 25, 23]. The most commonly used relational variables can be classified in different types. Some of them measure the central position of the firm in the network. This is the case of degree centrality, which shows the number of connections a firm has, or betweenness centrality, which measures the position of the firm as an intermediate between other firms [13].

Some variables representing homophily have also been used to detect fraudulent behavior. For example, [20] found that being directly connected or having board members connected to low tax paying firms increases the probability of being a low tax paying firm. [25] studied the characteristics of firms committing social security fraud in Belgium and found that interconnection between firms, belonging to a similar sector and geographical location, influenced their common behavior.

Another important group of network-related features represents the contagious effect. This means that those firms with strong relationships with fraudulent firms are prone to show fraudulent behavior as well. Some authors have represented this effect through the number of "risky" firms connected to a particular firm [20, 25]. An exposure score, built from a modification of the Personalized PageRank algorithm, has also been successfully used [24]. Specific graph structures in the firm neighborhood, such as triangles or quadrangles which include firms with certain characteristics, are also a way to represent the contagious effect [20, 25].

This paper aims to develop a model that represents the features of firms with VAT declaration mismatching. Mismatches show that at least one of the two firms filing tax returns has made an incorrect declaration, but not necessarily both. This characteristic differentiates this phenomenon from the other financial fraud cases analyzed above. In fact, few papers have specifically dealt with VAT declaration mismatching. Among them, [1] proposed an iterative method to identify the trustworthiness of firms in the complex network of VAT declarations. As was observed in other fraudulent activities, homophily is detected in the transaction network, i.e. firms tend to connect with other firms of similar trustworthiness. Other papers have applied techniques to identify suspicious structures of person-company relationships [23] and circular trading [21].

3. Material and Methods

3.1. Data

The source of the data used was the official administrative declaration of businesses liable for VAT concerning all transactions with persons or companies registered in the Canary Islands (Spain) in 2002 whose total amount exceeded 3,005.06 euros. The taxpayer has to declare independently the purchases and sales of all goods and services. Given the personal character of the data, we had no access to any information that could help to identify buyers or sellers and we were not allowed to work with the anonymized data outside the tax administration offices.

The main difficulty encountered treating the data was related to the identification of the activity developed by the tax filers. The codes used to classify the activities do not correspond to a standard economic classification of activities, and a correspondence between the classification made by the administrative source and the official nomenclature for activities (Spanish National Classification of Economic Activities - CNAE in its acronym in Spanish) had to be established. In addition, in many of the declarations the administrative classification of the activity was missing. This information had to be obtained from other administrative sources where the activity had been recorded. This information is not fully homogeneous because many taxpayers declare various secondary activities. The final assignment of principal activity was based on identifying the activity named as such in most statements.

The information that we considered had been separated at source into two files, one that records sales (551,721 operations) and another purchases (405,770 operations). In principle, each operation should be declared twice, once by the seller and once by the buyer, but some counterparts (e.g. individual buyers) are not obliged to declare VAT. A new file was prepared in which the transactions of each pair of agents was registered together with the amounts declared by the buyer and by the seller. The database was filtered to eliminate those transactions in which both counterparts coincide (311 operations). In addition, all transactions with no counterpart were also removed since, as mentioned, some counterparts do not have to declare VAT and we cannot therefore assign error labels to this kind of mismatch. After filtering, the final number of operations to be analysed amounted to 197,791 and included 32,886 firms. As the number of activity categories (57) was very high, it was reduced to 30.

3.2. Error types

Here we propose a metric for a firm's error in its VAT declarations, taking into account the type of mistake committed, which depends on the counterpart's declarations.

We follow a variation of the metric proposed by [1]. Using the same notation, we define $\alpha_a(a \to b)$ the amount sold by a to b, as declared by the seller a, and $\alpha_b(a \to b)$ the amount sold by a to b, as declared by the buyer b. We define two differences:

$$M(a,b) = \alpha_a(a \to b) - \alpha_b(a \to b), \text{ if } \alpha_a(a \to b) \geq \alpha_b(a \to b)$$
$$N(a,b) = -\alpha_a(a \to b) + \alpha_b(a \to b), \text{ if } \alpha_a(a \to b) \leq \alpha_b(a \to b)$$

Given the sum of the total amount declared by a and b

$$\Psi(a,b) = \alpha_a(a \to b) + \alpha_b(a \to b),$$

we define four types of weight in the operations:

$$T^{S^+}(a,b) = 1 - \frac{M(a,b)}{\Psi(a,b)}$$
$$T^{B^-}(b,a) = 1 - \frac{M(a,b)}{\Psi(a,b)}$$
$$T^{S^-}(a,b) = 1 - \frac{N(a,b)}{\Psi(a,b)}$$
$$T^{B^+}(b,a) = 1 - \frac{N(a,b)}{\Psi(a,b)}$$

Each agent is assigned an index between 0 and 1 indicating the level of correctness of each type in its operations:

$$T^i(a) = \frac{1}{\Upsilon(a)} \sum_{b \in \Upsilon(a)} T^i(a,b),$$

with $i = \{S^+, B^-, S^-, B^+\}$ and where $\Upsilon(a)$ is the set of neighbors (business partners) of a.

Therefore, T^{S^+} and T^{S^-} measure the level of correctness in VAT declarations when a works as a seller, T^{S^+} includes only upward errors and T^{S^-} only downward errors. Analogously, T^{B^+} and T^{B^-} measure the level of correctness in VAT declarations when a works as a buyer. T^{B^+} includes only upward errors and T^{B^-} only downward errors. Each firm is assigned a level of correctness of each type. Those firms without transactions corresponding to a specific type (e.g. firms that declare only sales have no value for T^{B^+} and T^{B^-}) are assigned the median value of the level of correctness of this type.

We define a firm's error of each type by re-coding the level of correctness to a binary response variable following the criteria:

$$E^i(a) = \begin{cases} 1 & T^i(a) < Q_1(T^i) - 1.5 \cdot IQR(T^i) \\ 0 & T^i(a) \geq Q_1(T^i) - 1.5 \cdot IQR(T^i) \end{cases} \quad i = \{S+, B-, S-, B+\}$$

where $Q_1(T^i)$ and $IQR(T^i)$ are the first quantile and the interquartile range ($Q_3(T^i) - Q_1(T^i)$) of the level of correctness T^i, respectively. Therefore, a firm is labeled as a type-i correct tax filer (CD-i) if its level of correctness T^i is above the first quantile minus one and a half the interquantile range ($E^i(a) = 0$), and as a type-i incorrect tax filer (ID-i) otherwise. This dichotomization criterion allows a better performance of the error classification algorithm.

3.3. Algorithm for error detection

In this study, an RF algorithm [7] was applied to determine the internal (firm level) and network features influencing the classification of a firm as ID-i, with $i = \{S^+, B^-, S^-, B^+\}$. The features considered and their description are shown in Table 1. The RF was designed using 70% of the data as training set, with 500 trees and taking the root square of the number of predictors as the number of variables randomly sampled as candidates at each split. Due to the high skewness of classes (CD includes at least 85% of firms for each error type), a balancing algorithm was applied.

We applied two models. Model 1 considers only the firm's features v_f, namely the branch in

Table 1
Variables used in the RF model.

Name	Description	Type
Firm features v_f		
branch	branch in which the firm is classified	categorical
turnover	firm's turnover	numeric
Network features v_n		
out.degree	firm's out-degree (number of buyers)	numeric
in.degree	firm's in-degree (number of sellers)	numeric
bet	firm's betweenness	numeric
bet.weights	firm's weighted betweenness	numeric
mean.buyer.i	Average T^i-correctness among the firm's buyers[1]	numeric
mean.seller.i	Average T^i-correctness among the firm's sellers[1]	numeric

[1] $i = \{S^+, B^-, S^-, B^+\}$

which the company is classified according to the CNAE and the sum of all transactions carried out by the firm and the share of the firm in its branch. Model 2 includes variables that depend on the network topology. Following previous studies, we included the feature related to the number of buyers/sellers of the firm (out/in-degree) and firm betweenness, which estimates the level of inter-mediation among all transactions in the business network. The weighted betweenness takes into account the transaction amount in the betweenness estimation [6]. Finally, in order to include potential contagious effects, the average T^i-correctness among the firm's sellers/buyers is also included.

These two models were applied to study the characteristics of ID-i firms, for each error type i as described in section 3.2. To evaluate the performance of the models, receiver operating characteristics (ROC) curves and the area under the curve (AUC) were used following [19], which measures the percentage of firms well classified by the model. The higher the AUC score, the better the model classifies the data. Additionally, we also show specificity (the proportion of the CD-i firms that are correctly classified) and sensitivity (the proportion of the ID-i firms that are correctly identified as such). The values for specificity and sensitivity are calculated by maximizing the F1 score. Models 1 and 2 were compared for each error type.

Additionally, we conducted a feature importance analysis in order to know which features most

Table 2
AUC for the four error types and two models.

	Error type			
	S^+	B^-	S^-	B^+
Model 1	0.561	0.516	0.555	0.563
Model 2	0.814	0.786	0.823	0.772

affect each error. We followed the method proposed by [10] and [18], where feature importance is estimated through the difference between the original loss function and the loss function using data resampled with the feature.

In combination with feature importance, we analyzed the marginal effect of these features in such probability. In models based on machine learning, such as RF, the marginal analysis can be performed using the partial dependence plot (PDP) introduced by [12]. In this paper we used the approach given by [15] in the *Pdp R package*. The marginal analysis shows the relationship between changes in the variable of interest and the output.

4. Results

Figure 2 shows the ROC curve for each error type i, $i = \{S^+, B^-, S^-, B^+\}$. As can be observed, when considering exclusively the firm's internal variables, the model behaves little better than a random predictor. However, when including the network features, the model substantially improves AUC for the four error types. The graphs show that the worst performance is obtained with type S^-, which describes the seller's errors when declaring a lower amount than reported in the buyer's declaration.

The AUC, sensitivity and specificity results confirm the previous observations (Table 2). The AUC shows up to a 50% increase when including network features. From an AUC of around 0.50 when assuming only firm features (almost a random predictor), it reaches maximum values of 0.77 and 0.78 when including network features for all error types, while for S^- AUC increases to 0.64.

The performance of Model 2 is further analysed in Table 3. As can be observed in the specificity results, a high percentage of firms classified as CD of any type are rightly assigned to this group.

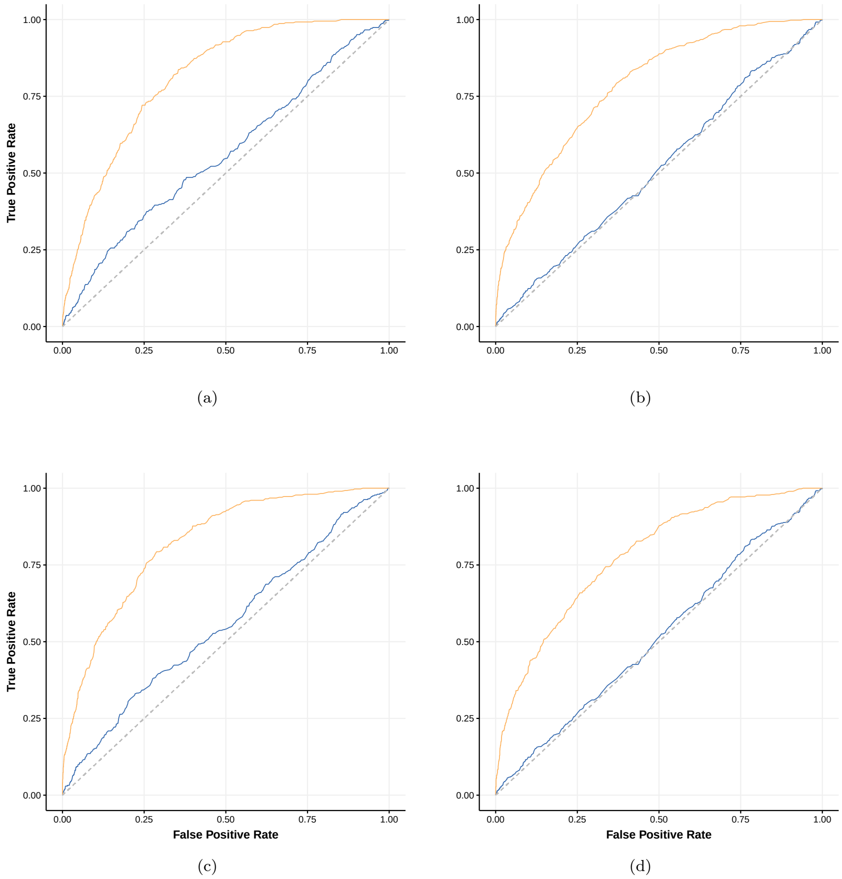

Figure 2: ROC Curves for the four error types. Model 1: blue curve; Model 2: orange curve; (a) Error S^{+}, (b) Error B^{-}, (c) Error S^{-}, (d) Error B^{+}.

Table 3
Performance results of model 2.

	Error type			
	S^{+}	B^{-}	S^{-}	B^{+}
Specificity	0.976	0.949	0.966	0.967
Sensitivity	0.157	0.317	0.256	0.188

However, this is not the case with firms classified as ID. If Tax Agencies order inspections according to this classification, the results show that a low percentage of correct tax filers are not well classified and therefore could be subject to unnecessary inspections. On the other hand, the model is able to identify almost 16% of incorrect tax filers.

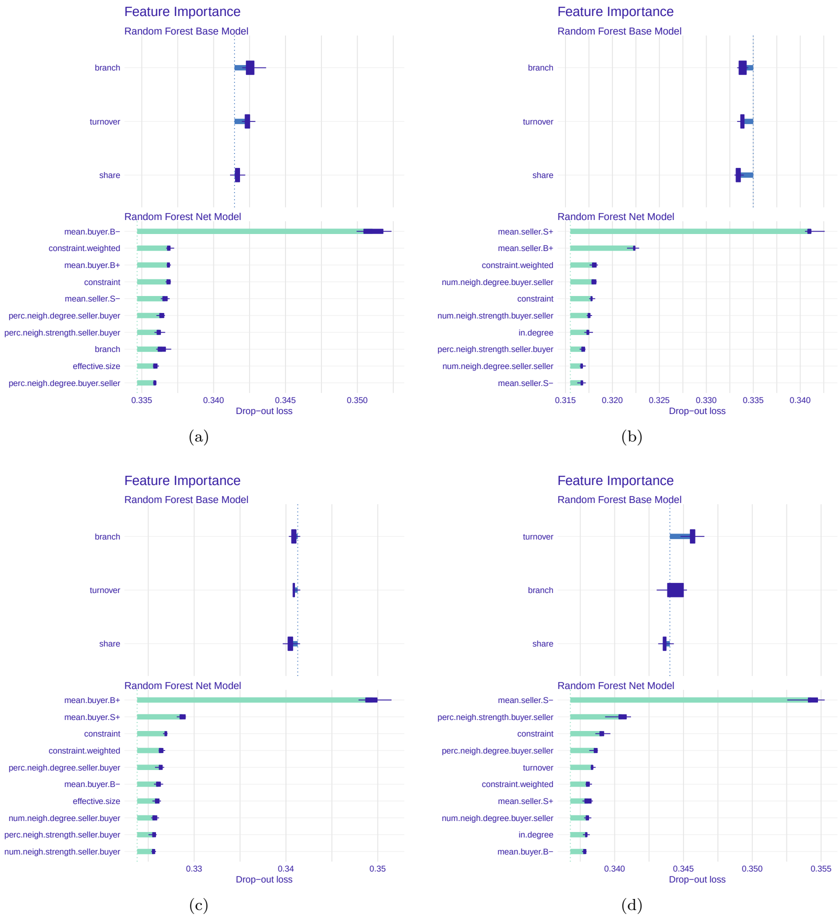

Figure 3: Feature importance plot. Model 1 (blue bars); Model 2 (green bars); (a) Error S^+, (b) Error B^-, (c) Error S^-, (d) Error B^+.

Figure 3 shows the results of the feature importance analysis. The end of the bar in each graph represents the value of the loss function once the specific feature is dropped and the beginning of the bar represents the value of the loss function of the model when all features are used. That is, feature importance in the classification algorithm can be depicted by bar length.

Since Model 1 is not useful to predict errors, the removal of any feature has little influence on model performance for the four error types. This result is expected. The results are more interesting for Model 2. The most important feature for each error type is average T^i-correctness among the firm's counterparts, pointing to a contagious effect among firms that file VAT returns. Specifically, the classification ID-S^+ (sellers declaring substantially higher amounts than buyers) depends on the amount of ID among its buyers, when declaring both more and less than its counterparts. The correspondence of ID-S^+ and ID-B^- is expected, since both are defined on the basis of the same operation (an operation with S^+-error for sellers corresponds to B^--error for buyers). The analysis shows that the wrong behavior of buyers in the other direction (buyers declaring substantially higher amounts than sellers) also influences sellers declaring substantially higher amounts than buyers.

A similar result occurs when taking the buyer's point of view. The classification of ID-B^+ mainly depends on the average T^{S^+}-correctness of its sellers. Again, buyers are more prone to declare higher amounts than their counterparts if their sellers are more prone to declare higher amounts that their counterparts. Therefore, the results show that the trend to declare a higher amount than the partner is contagious and independent of the role occupied (buyer or seller). In principle, this does not correspond to any intention of fraud, at least from the point of view of the seller, since taxes increase with the amount declared.

The factors influencing the S^--error are more varied. This type of error is mostly affected by neighbors' error of any type, together with the other network features included. Nevertheless, the performance of the model for error type S^- is not as positive as the rest of models (see Table 2).

The remaining network features occupy the first positions in feature importance, although with a significantly lower effect than that of neighbors. The most relevant factors are the number of sellers and buyers (out and in-degree) and the firm's betweenness. The internal firm characteristics exert only a small effect on the classification of each type of error. Turnover and to a lower extent branch only appear as influencing the classification of ID-S^+ firms.

Figure 4 presents the marginal effect of the most relevant features in Model 2 for each error type. Changes in the *in.degree* feature do not substantially change the average prediction of the model. Regarding the features *mean.buyer.B+*, *mean.buyer.B-* and *mean.buyer.S+*, the graphs show that an increase in the values of these features, that is the neighbors of the firm behaving correctly, causes the probability of the company being CD-i to be higher.

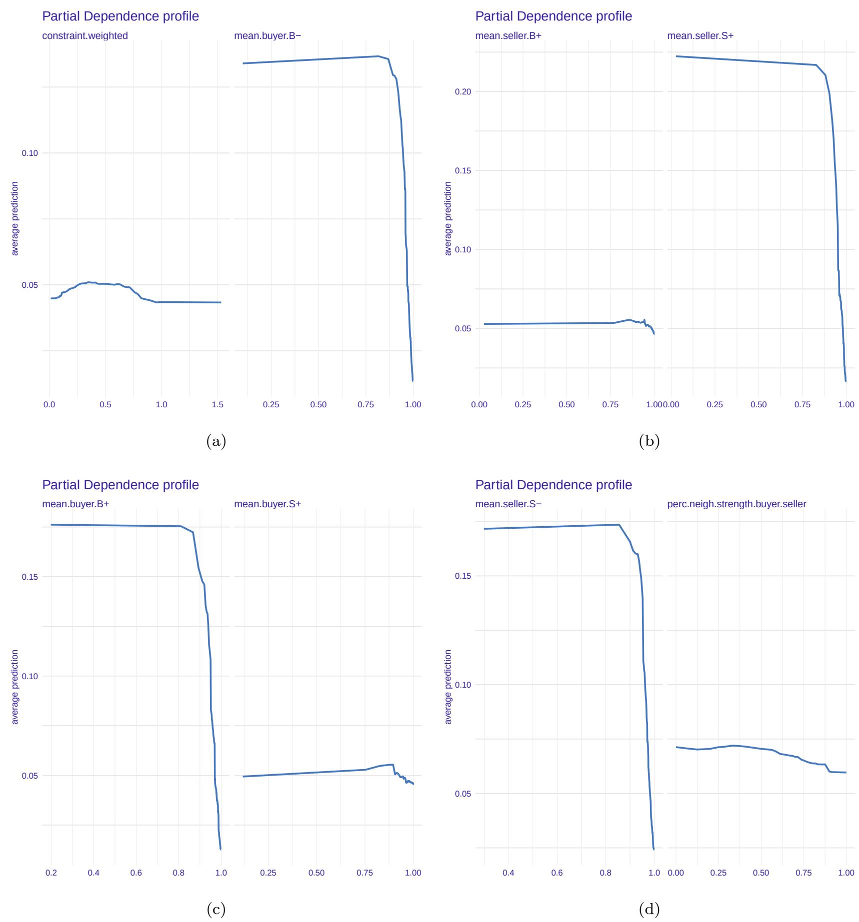

Figure 4: Partial dependency plots for the four types of error; (a) Error S^+, (b) Error B^-, (c) Error S^-, (d) Error B^+.

5. Conclusions

This paper provides a computational tool to identify the characteristics of firms with VAT declaration mismatches. Unlike other previous algorithms for fraud detection, the proposed method does not classify fraudulent and non-fraudulent firms, but rather firms that declare more or less than their counterparts when acting as sellers and buyers. The algorithm makes use of the transaction network formed by all firms and the commercial operations between them in an economy. This is a complex weighted and directed network and the method identifies the individual and network structural regularities of those firms that file incorrect VAT returns.

More specifically, the method was tested on the transaction network in a Spanish region in

2002. The findings show that the internal firm characteristics (branch, turnover and market share) do not help to predetermine whether the firm commits VAT declaration errors or not. However, a contagious effect among businesses was detected: those firms with VAT declaration errors in a specific direction (e.g. sellers declaring higher amounts than their counterparts) are usually related with other firms with VAT declaration errors in the same direction (e.g. buyers declaring higher amounts than their counterparts). This result points to a relational structure among firms that file incorrect VAT returns.

The paper presents several limitations. One of them is the data source, since the method was applied to a transaction network in a single year. An inter-temporal validation would be more informative about its ability to detect mistaken tax filers. Nevertheless, the model presented in this paper is novel in its ability to detect suspicious VAT declaration structures and can be applied to other regions where VAT is applied. Moreover, this research confirms the need for tax agencies to incorporate complex network metrics in the process of identifying taxpayer misbehavior.

References

[1] Mauro Barone and Michele Coscia. Birds of a feather scam together : Trustworthiness homophily in a business network. *Social Networks*, 54:228–237, 2018.

[2] Siddhartha Bhattacharyya, Sanjeev Jha, Kurian Tharakunnel, and J. Christopher Westland. Data mining for credit card fraud: A comparative study. *Decision Support Systems*, 50(3):602–613, 2011.

[3] Kay Blaufus, Jonathan Bob, Philipp E Otto, Nadja Wolf, Kay Blaufus, Jonathan Bob, Philipp E Otto, Nadja Wolf, The Effect, and K A Y Blaufus. The Effect of Tax Privacy on Tax Compliance – An Experimental Investigation The Effect of Tax Privacy on Tax Compliance – An Experimental Investigation. *European Accounting Review*, 26(3):561–580, 2017.

[4] Stefano Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308, 2006.

[5] Aurélie Bonein. La réciprocité , entre psychologie et rationalité économique. *Revue Française d'économie*, 23:203–232, 2018.

[6] Ulrik Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.

[7] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, oct 2001.

[8] Roberta Calvet and James Alm. Empathy , sympathy , and tax compliance. *Journal of Economic Psychology*, 40:62–82, 2014.

[9] Pamela Castellón González and Juan D. Velásquez. Characterization and detection of taxpayers with false invoices using data mining techniques. *Expert Systems with Applications*, 40(5):1427–1436, 2013.

[10] A Fisher and C Rudin. Model class reliance: Variable importance measures for any machine learning model class, from the" rashomon" perspective. *ArXiv:1801.01489, F Dominici - arXiv preprint*, 2018.

[11] Bruno S Frey and Benno Torgler. Tax morale and conditional cooperation. *Journal of Comparative Economics*, 35:136–159, 2007.

[12] JH Friedman. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232, 2001.

[13] Andrea Fronzetti Colladon and Elisa Remondi. Using social network analysis to prevent money laundering. *Expert Systems with Applications*, 67:49–58, 2017.

[14] Herbert Gintis. Strong Reciprocity and Human Sociality. *Journal of Theoretical Biiology*, 206:169–179, 2000.

[15] Brandon M Greenwell. pdp: An R Package for Constructing Partial Dependence Plots. *The R Journal*, 9(1):421–436, 2017.

[16] M Keen and S Smith. VAT Fraud and Evasion: What Do We Know and What Can Be Done? *National Tax Journal*, LIX(4):861–887, 2006.

[17] Michael Keen. VAT attacks ! *International Tax and Public Finance*, 14:365–381, 2007.

[18] Przemyss Law Biecek. DALEX: Explainers for Complex Predictive Models in R. Technical report, 2018.

[19] Charles X. Ling, Jin Huang, and Harry Zhang. AUC: A statistically consistent and more discriminating measure than accuracy. *IJCAI International Joint Conference on Artificial Intelligence*, pages 519–524, 2003.

[20] Jasmien Lismont, Eddy Cardinaels, Liesbeth Bruynseels, Sander De Groote, Bart Baesens, Wilfried Lemahieu, and Jan Vanthienen. Predicting tax avoidance by means of social network analytics. *Decision Support Systems*, 108:13–24, 2018.

[21] Priya Mehta, Jithin Mathews, Sandeep Kumar, K. Suryamukhi, Ch Sobhan Babu, and S. V. Kasi Visweswara Rao. Big data analytics for nabbing fraudulent transactions in taxation system. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11514 LNCS:95–109, 2019.

[22] E. W.T. Ngai, Yong Hu, Y. H. Wong, Yijun Chen, and Xin Sun. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3):559–569, 2011.

[23] Feng Tian, Tian Lan, Kuo Ming Chao, Nick Godwin, Qinghua Zheng, Nazaraf Shah, and Fan Zhang. Mining suspicious tax evasion groups in big data. *IEEE Transactions on Knowledge and Data Engineering*, 28(10):2651–2664, 2016.

[24] Véronique Van Vlasselaer, Tina Eliassi-Rad, Leman Akoglu, Monique Snoeck, and Bart Baesens. AFRAID: Fraud detection via active inference in time-evolving social networks. *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2015*, (1):659–666, 2015.

[25] Véronique Van Vlasselaer, Tina Eliassi-Rad, Leman Akoglu, Monique Snoeck, and Bart Baesens. GOTCHA! Network-based fraud detection for social security fraud. *Management Science*, 63(9):3090–3110, 2017.

[26] Jarrod West and Maumita Bhattacharya. Intelligent financial fraud detection: A comprehensive review. *Computers and Security*, 57:47–66, 2016.

[27] M. Zanin, D. Papo, P. A. Sousa, E. Menasalvas, A. Nicchi, E. Kubik, and S. Boccaletti. Combining complex networks and data mining: Why and how. *Physics Reports*, 635:1–44, 2016.