Decision Support Systems 141 (2021) 113464

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Identifying business misreporting in VAT using network analysis

Christian González-Martel [a], Juan M. Hernández [a,b,*], Casiano 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 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 error were assigned to each business in the network, defined from the mismatch between the amount declared by the firm in question and its counterpart. We applied a random forest algorithm to detect which firm-related and which network-related 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[1]

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 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 [1]. 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 [2].

To illustrate how the system works, Fig. 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 factors in a VAT declaration network that best characterize tax statement mismatches. Some of the studied factors correspond to the characteristics of the firm in question (branch of economic activity, size, etc.), but so-called relational factors also play an essential role in the characterization of mismatches. For example, contagious effects (firms mimic the behavior of counterparts) 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 [3] provides useful metrics to

* Corresponding author.
E-mail address: juan.hernandez@ulpgc.es (J.M. Hernández).
[1] ORCID(s): 0000-0003-1081-0843 (Christian González-Martel); 0000-0001-6897-5179 (Juan M Hernández)*Abbrevations:* VAT: Value-added tax; RF: Random forest; LR: Logistic regression

https://doi.org/10.1016/j.dss.2020.113464
Received 18 February 2020; Received in revised form 24 November 2020; Accepted 25 November 2020
Available online 10 December 2020

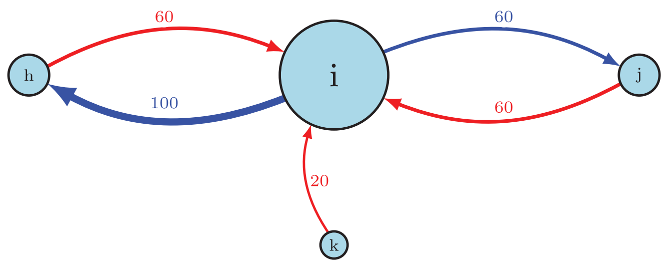

Fig. 1. Example of VAT declaration network. Balls *i*, *j*, *k* and *h* represent firms, arrows represent sale and purchase declarations.

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 [4].

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 emergent structure of firms filing incorrect VAT returns? 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-related characteristics that explain errors in VAT declarations.
- We provide a model to estimate firm- and network-related factors 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 [5,6]. Among them, classification algorithms are commonly used. These methods try to identify features that describe suspicious firms, and start from a predefined 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 [7].

Traditionally, the models used for fraud detection take into account exclusively local (intrinsic) factors, which refer to characteristics of the firm [8,9] 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, [10] 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 iniquitous/disadvantageous to them. This result is contrary to the predictions of a self-interest model [11], where agents follow in the footsteps of potential future material gains. [12] 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. [13] 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 earlier periods of the experiment whereas the contagion effect dominates in the latter periods and overrides the shame effect. In other words, knowing what others do strongly influences agents' attitudes towards tax evasion.

[14] 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 incorporated the role of relational factors in the models, obtaining more accurate results than when using only local variables [e.g. 15, 16, 17, 8, 18]. 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 intermediacy of the firm in relation to other firms [16].

Some variables representing homophily have also been used to detect fraudulent behavior. For example, [15] 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. [8] 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 factors 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 [8,15]. An exposure score, built from a modification of the Personalized PageRank algorithm, has also been successfully used [17]. 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 [8,15].

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, [19] 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 [18] and circular trading [20].

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 3005.06 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 by this administrative body 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 in question 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.

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. individuals who are 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 the buyer was also the seller (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 analyzed amounted to 197,767 and included 32,886 firms. As the number of activity categories (57) was very high, it was reduced to 30.

Fig. 2 (left) presents the in-degree (number of sellers) and out-degree (number of buyers) distribution of the VAT declaration network. In general, sellers have more connections than buyers in the network, as revealed by the heavier tail of the out-degree distribution. In fact, following the procedure in [21], the out-degree fits a power-law distribution $p(k) \sim k^{\alpha}$, with $\alpha = 2.62$ starting from degree $k_{min} = 44$.

A node's strength is defined as the sum of the amounts declared by a firm, when working as a seller or as a buyer. Then, the right-hand side of Fig. 2 presents out-strength and in-strength distribution, according to the amount declared by sellers and buyers, respectively. As can be observed, the two distributions are quite similar, presenting a very heavy tail. This characteristic reveals that there are a few firms who trade a much larger amount than the rest.

In order to describe in greater depth the characteristics of the VAT transaction network, Table 1 presents some network-level metrics [22,23] of the empirical network and compares them with those obtained with a simulated network with a similar degree distribution. We used the R package *igraph* [24] to calculate some of these metrics and other network variables. For the sake of simplicity, we do not differentiate when a firm works as seller or buyer and therefore the network is undirected. In this case, the network fits a power law well, with $\alpha = 3.04$ and $k_{min} = 88$.

The data in Table 1 reveals that VAT transactions correspond to a sparse network, showing a small-world effect, as is usually the case in scale-free distributions. In general, the centralization metric adopts values between 0 and 1, indicating how close the network is to a star-like

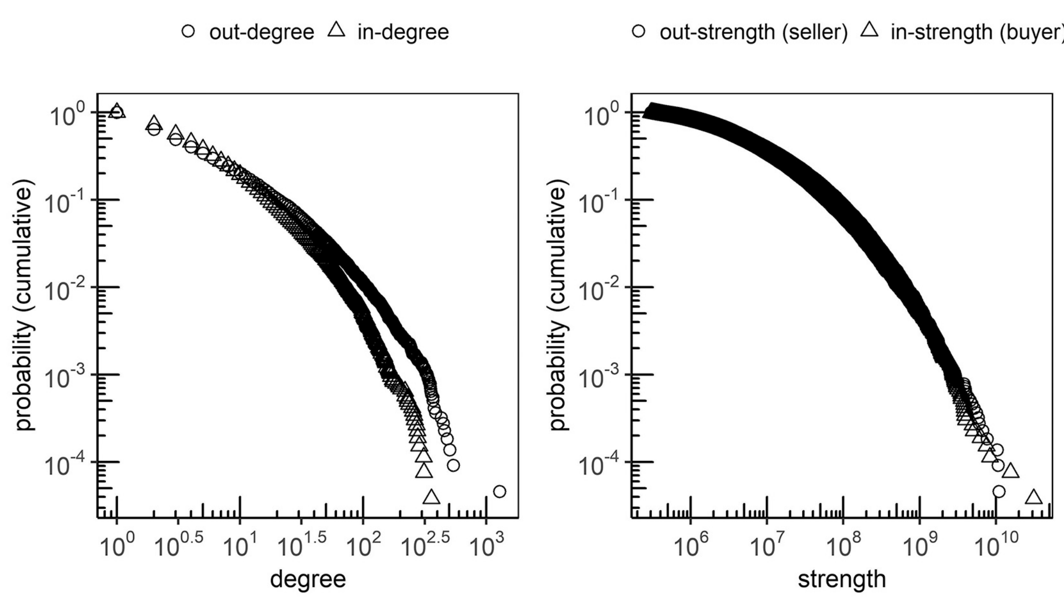

Fig. 2. VAT declaration network: out- and in-degree distribution (left); out-strength (as declared by sellers) and in-strength (as declared by buyers) (right).

Table 1
Network metrics for the VAT declaration network and a simulated power-law distributed network with the same number of nodes, edges and identical parameters.

	sample	simulated power-law
density	0.00037	0.00037
α	3.04	3.04
k_{min}	88	0
Average distance	4.09668	3.90064
Clustering coefficient	0.03093	0.00212
Centralization	0.04121	0.03105
Assortativity (degree)	−0.03425	−0.00576

network. The low value of this metric in Table 1 shows that transactions are not centralized around a few firms, as is also expected in a random power-law network. However, the clustering coefficient in the real network is substantially higher than in the simulated power-law network. This means that the VAT declaration network has more triangular relationships than would be expected by chance. The assortativity index adopts values between −1 and 1 and shows the trend of a firm to trade with firms with similar degree (positive value) or dissimilar degree (negative value). The values are close to 0, revealing that there is no marked tendency.

3.2. Error types

Here, we propose a metric for a firm's error in its VAT declarations, taking into account the type of error committed, which depends on the counterpart's declarations. In order to avoid scale effects, we take normalized values of the firm's errors.

We follow a variation of the metric proposed by [19]. 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 weights 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 of error 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 (counterparts in transactions) 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 quartile 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 quartile minus one and a half the interquartile 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 [25] was applied to determine the internal (firm level) and network features influencing the classification of a firm as ID-i, with $i = \{S^+, B^-, S^-, B^+\}$. An alternative methodology is logistic regression (LR). However, we opted to adopt RF as this methodology tends to outperform the LR approach [26,27]. Nevertheless, we show the prediction performance results using both methodologies for the sake of comparison. The factors considered and their descriptions are shown in Table 2. 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 apply two models. Model 1 considers only the firm's features, v_f, namely the branch of economic activity in which the company is classified according to the CNAE, the sum of all transactions carried out by the firm, and the market share of the firm in that economic activity. Model 2 incorporates additional features that depend on the topology of the network. Following previous studies [8,15,16], we include features representing the firm's characteristics, including the number of buyers/sellers (out/in-degree) and the out- and in-strength of sellers and buyers, respectively. We also include some centrality metrics, such as the alpha centrality [28], where relevant firms are those surrounded by relevant firms. It is related to eigenvector and Page Rank centrality [22] and can be calculated disregarding or taking into account the weight (transaction amount) of the links.

Additionally, we consider some features representing the influence of first-order network neighbors (immediate neighbors) on the behavior of a firm with such neighbors. Specifically, we adopt Burt's effective size and constraint [29,30]. Effective size is the firm's degree minus the average degree of its neighbors, showing the firm's size with respect to its neighbors. Constraint measures the firm's capacity to sell to firms who do not trade between each other. It is therefore a metric of the level of influence of the firm over its buyers. Constraint can be calculated for weighted and unweighted networks.

In order to include potential contagious effects, the average T^i-correctness among the firm's sellers/buyers is also considered. By means of these features, we characterize contagion among neighbors when playing complementary roles in the transaction network (seller and buyer) and when playing the same role. For instance, the latter would be represented by the contagion effect of a buyer to its seller when the former acts as seller with third parties. In addition, the possible mimic of the behavior of neighbors who are relevant is represented by the number/percentage of the firm's neighbors with a larger degree/strength. The homophily effects are also tested by means of the number/percentage of the neighbors of a firm belonging to the same branch.

Table 2
Features used in the RF model.

Name	Description	Type
Firm features v_f		
branch	branch of economic activity in which the firm is classified	categorical
turnover	firm's turnover	numeric
share	firm's market share in its branch of economic activity	numeric
Network features v_n		
Individual firm level		
out.degree	firm's out-degree (number of buyers)	numeric
in.degree	firm's in-degree (number of sellers)	numeric
out.strength	firm's out-strength (traded amount of sellers)	numeric
in.strength	firm's in-strength (traded amount of buyers)	numeric
alpha.centrality	firm's alpha centrality	numeric
alpha.centrality.weighted	firm's weighted alpha centrality	numeric
First-order neighbors level		
effective.size	firm's effective size	numeric
constraint	firm's constraint	numeric
constraint.weighted	firm's weighted constraint	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*	numeric
num.neigh.degree.j.h	number of neighbors h of firm j with larger degree[2]	numeric
perc.neigh.degree.j.h	percentage of neighbors h of firm j with larger degree	numeric
num.neigh.strength.j.h	number of neighbors h of firm j with larger strength	numeric
perc.neigh.strength.j.h	percentage of neighbors h of firm j with larger strength	numeric
num.neigh.branch.j.h	number of neighbors h of firm j belonging to the same branch	numeric
perc.neigh.branch.j.h	percentage of neighbors h of firm j belonging to the same branch	numeric
Network level		
betweenness	firm's betweenness	numeric
betweenness.weighted	firm's weighted betweenness	numeric

[1] $i = \{S^+, B^-, S^-, B^+\}$.
[2] The firm works as a $j = \{seller, buyer\}$ and the neighbor works as a $h = \{seller, buyer\}$.

Finally, we also consider some features representing the firm's position in the network as a whole. More specifically, firm betweenness estimates the level of intermediacy among all transactions in the business network, while weighted betweenness takes into account the transaction amount in the betweenness estimation [31].

The two models were applied to study the characteristics of ID-i firms, for each error type i 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 [32], 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 F-1 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 affect each error. We followed the method proposed by [33,34], 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 the features on the probability of a firm making incorrect VAT declarations. In models based on machine learning, such as RF, marginal analyses can be performed using the partial dependence plot (PDP) introduced by [35]. In this paper, we used the approach given by [36] in the R package *Pdp*. The marginal analysis shows the relationship between changes in the feature of interest and the output.

4. Results

Fig. 3 shows the ROC curve for each error type i, $i = \{S^+, B^-, S^-, B^+\}$. To compare the fit results, we include the results for Model 2 using RF and LR. As can be observed, when considering exclusively a firm's internal factors, the model behaves little better than a random predictor. However, when including the network factors, the model substantially improves AUC for the four error types. In general, the ROC curve using LR in Model 2 is slightly below the ROC obtained with RF.

The AUC, sensitivity and specificity results confirm previous observations (Table 3). The AUC shows up to a 50% increase when including network factors. From an AUC of around 0.50 when assuming only firm factors (almost a random predictor), it reaches maximum values of 0.82 and 0.78 when including network factors for all error types. The AUC for Model 2 using LR is below the maximum values obtained with RF in all errors.

The performance of Model 2 is further analyzed in Table 4. 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. However, this is not the case with firms classified as ID. If tax agencies were to order inspections according to this classification, the results show that a small percentage of correct tax filers would not be correctly classified and therefore could be subject to unnecessary inspections. On the other hand, the model is able to identify almost 16% of S^+ incorrect tax filers.

Fig. 4 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, Fig. 4a) 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 errors are complementary, in the sense that an operation with S^+-error for a seller corresponds to B^--error for its buyer. Nevertheless, we have to bear in mind that the estimation of ID-S^+ and ID-B^- takes into account all operations each agent has with other firms. The remaining contagious effects do not correspond to the error definition. More specifically, buyers declaring substantially higher amounts than sellers (ID-B^+) make sellers declare substantially higher amounts with other buyers.

A similar result occurs when taking the buyer's point of view. The classification of ID-B^- (Fig. 4b) mainly depends on the average T^{S+}-correctness of its sellers. Again, another contagious effect from its sellers arises in the second place of importance: firms declaring a higher amount that its counterparts when acting as buyers (ID-B^+).

In order to interpret the results above, we have to take into account that errors ID-S^+ and ID-B^- do not correspond to any intention of fraud, since taxes increase with the amount declared by the seller. Therefore, in the case of ID-B^-, the results show that this kind of error arises as a consequence of the trend to declare in excess of their counterparts when acting as sellers and buyers.

The factors influencing the S^- and B^+ error are more varied. The most influential factor is the corresponding error from their counterparts

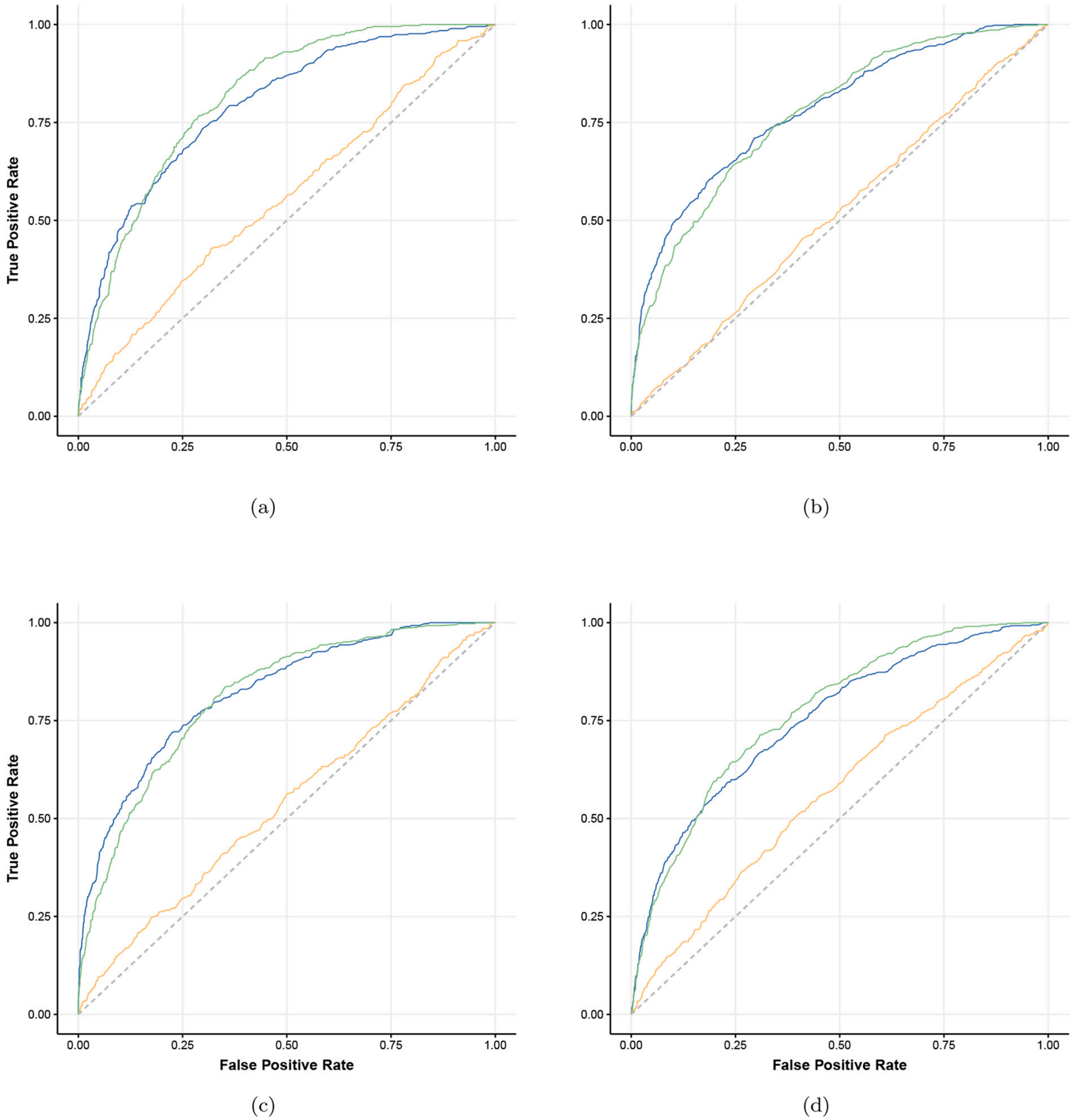

Fig. 3. ROC Curves for the four error types. Model 1: orange curve; Model 2: green curve; Model 2 (LR): blue curve; (a) Error S^{+}, (b) Error B^{-}, (c) Error S^{-}, (d) Error B^{+}. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
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	0.813	0.786	0.822	0.772
Model 2 (LR)	0.793	0.782	0.821	0.751

Table 4
Performance results of Model 2.

	Error type			
	S^{+}	B^{-}	S^{-}	B^{+}
Specificity	0.976	0.949	0.966	0.967
Sensitivity	0.157	0.300	0.256	0.188

(ID-B^{+} and ID-S^{-}, respectively). In addition, sellers with S^{-} error are influenced by buyers' behavior in the other direction when acting as sellers (S^{+}-error). This result suggests that firms do not act following an imitative, but rather an opposing behavior.

Some other network factors occupy the first positions in feature importance for the four types of error, although with a significantly lower effect than that of neighbors. It is noteworthy that constraint (weighted and unweighted) is a relevant metric to determine IDs. This factor was also obtained as a determinant of fraud behavior in previous studies [16] and can be interpreted in this context as the trend of VAT error for those firms trading with firms who do not trade between each other. Therefore, the level of social capital of a firm influences declaration mismatches.

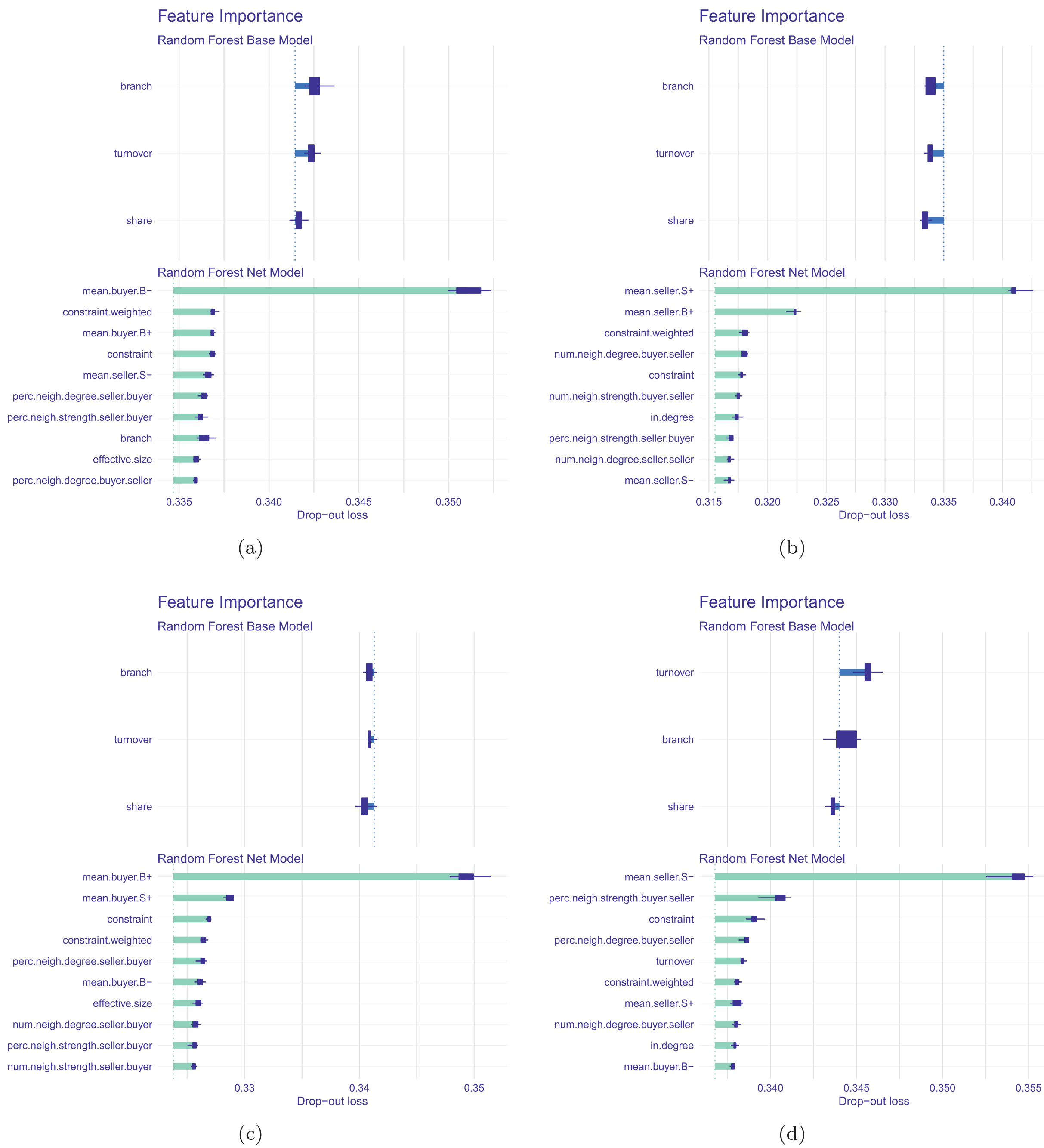

Fig. 4. Variable importance plot. Model 1 (blue bars); Model 2 (green bars); (a) Error S^+, (b) Error B^-, (c) Error S^-, (d) Error B^+. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The size of the counterparts also influences being ID firms of any type, as observed by the importance of the number and percentage of neighbors with a larger degree and strength, mostly for error B^+. Other relevant network variables are the number of sellers and buyers (out and in-degree), but only for ID-S^- and ID-B^+. However, internal firm characteristics exert only a small effect on the classification of each type of error. Branch of economic activity only appears as influencing the classification of ID-S^+ firms. No homophily effects of firms belonging to the same branch is observed. Network level factors, represented by the firm's betweenness, are also irrelevant for determining any type of error.

The results of Model 2 using the LR approach can be found in the Appendix. The endogenous variable is ID-*i* and the features of Table 2 are the explanatory variables. In general, the significant factors in the regression model correspond to those occupying the first positions in Fig. 4. There are some factors, such as out- and in-strength, that positively influence the trend to become an incorrect tax filer of some types of error (ID-B^- and ID-B^+) which were not captured by the RF model. Nevertheless, as was commented in section 3.3, we decided to adopt the estimations obtained with RF, since this method obtains a better fit to the data. Moreover, the RF methodology allows the capture of the non-linear effect of changes in the explanatory features, as is shown below.

Fig. 5 presents the marginal effect of the most relevant features in Model 2 for each error type. Each graph in the Figure shows how the feature influences the probability of being ID-*i* depending on the range of values of the feature. Observe that some features representing the counterpart's behavior (*mean.buyer.B*+, *mean.seller.S*+, *mean.buyer.S*+ and *mean.seller.S*-) have similar effects in the different types of error. The graphs show that when a firm's neighbors behave incorrectly in a range

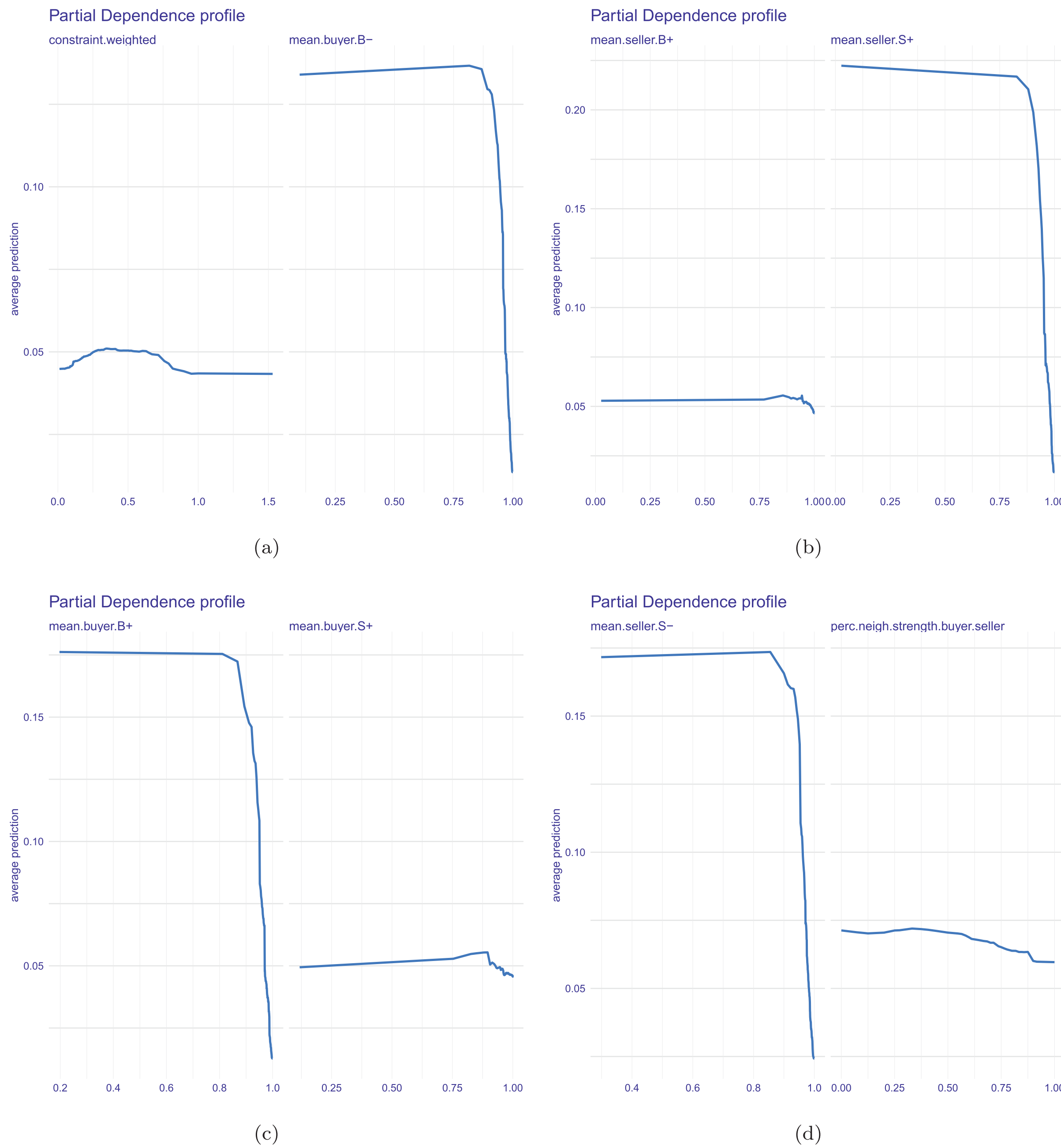

Fig. 5. Partial dependency plots for the four types of errors. (a) Error S^+, (b) Error B^-, (c) Error S^-, (d) Error B^+.

of values between 0 and 0.9 approximately, the probability of the company being ID-*i* is relatively constant. However, this probability dramatically decreases when neighbors behave closer to correct (range of values between 0.9 and 1). This effect is partially expected since the firm's and the counterpart's errors are complementary, although the counterpart's error include operations with third parties. A similar marginal effect on B^--error and S^--error can be found with *mean.seller.B*+ and *mean.buyer.S*+, respectively. These factors do not represent a complementary error, but also present two phases (the effect for values under 0.9 is higher than for values above 0.9), although not so extreme as in the other four features. These results reveal a contagious effect, in the sense that a firm has lower declaration errors when its counterparts have lower declaration errors in their operations with third parties.

The marginal effect of other network factors is noteworthy. Specifically, changes in the *constraint.weighted* feature positively influence the probability of the firm being ID-S^+, for low values of this feature (below 0.5 approximately), but turns negative for larger values (up to 1 approximately) and thereafter remains constant. This result implies that firms with high social capital (those who trade with other firms who do not trade with each other) are more prone to declare selling in excess (ID-S^+) than firms with less influence on their transaction network. Since selling in excess is not a fraudulent behavior, the result means that firms with high social capital are more prone to avoid fraudulent behavior than those with less social capital.

Finally, the negative effect of the percentage of neighbors with larger strength on the probability of being ID-B^+ shows the influence exerted by trading with smaller and larger firms in making correct declarations. The graph shows that the firm is more prone to declare purchases in excess when trading with smaller counterparts than when trading with large counterparts. This result points to a positive influence (in terms of correct declarations) of large firms on their counterparts.

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 characteristics of those firms that file incorrect VAT returns.

More specifically, the method was tested on the transaction network in a Spanish region in 2002. Three main findings can be extracted from the analysis. First, the internal firm characteristics (turnover and market share) do not help to predetermine whether the firm commits VAT declaration errors or not. Additionally, firms belonging to a particular branch are not more prone to follow a similar behavior. Second, a contagious effect among businesses was detected in the sense that a firm is more prone to declare correctly when their counterparts declare correctly not only with that firm but with other counterparts, acting as seller and buyer. Third, firms with high social capital (firms trading with other firms who do not trade between each other) are more prone to avoid fraudulent behavior, at the same time, large firms (in terms of high trading volume) also favor the correct declaration of their counterparts.

Several recommendations for tax agencies when deciding inspections can be derived from the findings of this study. When defining their fraud detection programs, they should not concentrate their efforts in relevant firms in terms of high social capital because they tend to declare correctly, selecting firms by branch would not be efficient neither. Fraud detection efforts should be oriented towards low and middle-size firms and, once fraud is detected, inspection should be extended to their counterparts.

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 the way it detects 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.

Acknowledgements

The authors wish to thank the editor and two anonymous reviewers for their comments which have improved the final quality of the manuscript. We also thank Stephen Burns for proof reading the paper. This work was supported by the by the Economy and Finance Council of the Regional Government of the Canary Islands, Spain [grant numbers FULP: CN45/08240/57/100]; the University of Las Palmas de Gran Canaria [grant number COVID 19–04]. We would like to specially thank the Councilor of Economy, Knowledge and Employment of the Regional Government of the Canary Islands (SPAIN) for allowing us to use the results of the previously mentioned project.

Declarations of interest

None.

Appendix A. Logistic regression results

For each error, we built a logistic regression model to determine ID-i, $i = \{S^+, B^-, S^-, B^+\}$. Tables A.1, A.2, A.3 and A.4 show the coefficients, standard errors, z-values, p-values and standardized coefficients of the significant variables of each model.

Table A1
Coefficients, standard errors, z-values, p-values and standardized coefficients for the logistic regression with ID-S^+.

Variable	coefficients	std_error	z_value	p_value	standardized_coeff.
Intercept	−25.717	0.931	−27.634	0.000	−0.871
mean.buyer.B-	−4.487	0.857	−5.236	0.000	−0.128
mean.seller.S-	−2.360	0.860	−2.743	0.006	−0.081
constraint	−0.710	0.232	−3.062	0.002	−0.240
mean.buyer.B+	−0.440	0.219	−2.010	0.044	−0.138
bet	−0.099	0.056	−1.764	0.078	−0.265
num.neigh.degree.seller.buyer	−0.031	0.018	−1.693	0.090	−0.171
perc.neigh.strength.seller.buyer	< 0.001	< 0.001	2.654	0.008	0.153
mean.seller.B+	0.019	0.011	1.746	0.081	0.207
branch.FF	0.047	0.018	2.575	0.010	0.193
perc.neigh.strength.buyer.buyer	0.378	0.223	1.699	0.089	0.121
num.neigh.degree.seller.seller	0.571	0.313	1.826	0.068	0.571
perc.neigh.strength.seller.buyer	1.839	0.917	2.003	0.045	0.069
num.neigh.strength.seller.buyer	28.620	1.825	15.684	0.000	−2.403

Table A2
Coefficients, standard errors, z-values, p-values and standardized coefficients for the logistic regression with ID-B^-.

Variable	coefficients	std_error	z_value	p_value	standardized_coeff.
Intercept	34.101	1.809	18.847	0.000	−2.32
mean.seller.S+	−33.862	0.991	−34.16	0.000	−1.166
mean.seller.B+	−2.481	0.6	−4.137	0.000	−0.105
mean.buyer.B-	−1.647	0.898	−1.835	0.067	−0.042
mean.seller.B-	1.478	0.872	1.694	0.090	0.047
constraint	−1.001	0.207	−4.824	0.000	−0.319
perc.neigh.strength.buyer.seller	0.419	0.252	1.663	0.096	0.115
perc.neigh.degree.buyer.buyer	−0.149	0.064	−2.303	0.021	−0.345

(*continued on next page*)

Table A2 (*continued*)

Variable	coefficients	std_error	z_value	p_value	standardized_coeff.
num.neigh.degree.buyer.seller	−0.066	0.023	−2.843	0.004	−0.347
effective.size	−0.044	0.023	−1.912	0.056	−1.053
out.degree	0.038	0.023	1.664	0.096	0.931
num.neigh.degree.buyer.buyer	0.03	0.009	3.274	0.001	0.305
betweenness	< 0.001	< 0.001	2.082	0.037	0.119
in.strength	< 0.001	< 0.001	2.781	0.005	0.119
out.strengh	< 0.001	< 0.001	1.649	0.099	0.078
turnover	< 0.001	< 0.001	−1.708	0.088	−0.111

Table A3
Coefficients, standard errors, z-values, p-values and standardized coefficients for the logistic regression with ID-S^{-}.

Variable	coefficients	std_error	z_value	p_value	standardized_coeff.
Intercept	32.27	1.961	16.457	0.000	−2.411
mean.buyer.B+	−29.342	0.992	−29.574	0.000	−1.273
mean.buyer.B-	−2.997	0.898	−3.337	0.001	−0.091
mean.buyer.S+	−1.447	0.562	−2.574	0.01	−0.076
branch.DN	0.783	0.341	2.294	0.022	0.783
constraint	−0.626	0.243	−2.58	0.01	−0.214
perc.neigh.degree.seller.seller	−0.184	0.054	−3.386	0.001	−0.897
perc.neigh.streng.seller.seller	0.157	0.054	2.909	0.004	0.693
num.neigh.degree.seller.buyer	0.048	0.021	2.323	0.02	0.202
num.neigh.degree.seller.seller	0.023	0.013	1.852	0.064	0.256

Table A4
Coefficients, standard errors, z-values, p-values and standardized coefficients for the logistic regression with ID-B^{+}.

Variable	coefficients	std_error	z_value	p_value	standardized_coeff.
Intercept	27.716	1.618	17.13	0.000	−2.343
mean.buyer.S-	−24.311	0.866	−28.079	0.000	−0.816
mean.buyer.B-	−4.002	0.775	−5.162	0.000	−0.106
mean.buyer.B+	−1.849	0.824	−2.244	0.025	−0.053
perc.neigh.strengh.buyer.seller	−1.031	0.213	−4.844	0.000	−0.283
branch.DG	0.789	0.431	1.829	0.067	0.789
perc.neigh.degree.buyer.seller	0.785	0.224	3.504	0.000	0.199
constraint	−0.628	0.194	−3.246	0.001	−0.199
branch.GG	−0.488	0.297	−1.646	0.100	−0.488
constraint.weighted	−0.438	0.172	−2.545	0.011	−0.139
num.neigh.degree.buyer.buyer	0.019	0.008	2.439	0.015	0.184
in.degree	0.011	0.005	2.368	0.018	0.162
betweenness	0.001	< 0.001	1.767	0.077	0.039
in.strength	< 0.001	< 0.001	−2.671	0.008	−0.149
turnover	< 0.001	< 0.001	3.114	0.002	0.375
out.strength	< 0.001	< 0.001	−2.237	0.025	−0.135
betweenness.weighted	< 0.001	< 0.001	5.694	0.000	42.845

References

[1] M. Keen, S. Smith, VAT fraud and evasion: What do we know and what can be done? National Tax Journal LIX (4) (2006) 861–887.
[2] M. Keen, VAT attacks !, Int. Tax Public Financ. 14 (2007) 365–381, https://doi.org/10.1007/s10797-007-9037-9.
[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.U. Hwang, Complex networks: structure and dynamics, Phys. Rep. 424 (4–5) (2006) 175–308, https://doi.org/10.1016/j.physrep.2005.10.009.
[4] M. Zanin, D. Papo, P.A. Sousa, E. Menasalvas, A. Nicchi, E. Kubik, S. Boccaletti, Combining complex networks and data mining: why and how, Phys. Rep. 635 (2016) 1–44, https://doi.org/10.1016/j.physrep.2016.04.005.
[5] E.W. Ngai, Y. Hu, Y.H. Wong, Y. Chen, X. Sun, The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature, Decis. Support. Syst. 50 (3) (2011) 559–569, https://doi.org/10.1016/j.dss.2010.08.006.
[6] J. West, M. Bhattacharya, Intelligent financial fraud detection: a comprehensive review, Computers and Security 57 (2016) 47–66, https://doi.org/10.1016/j.cose.2015.09.005.
[7] S. Bhattacharyya, S. Jha, K. Tharakunnel, J.C. Westland, Data mining for credit card fraud: a comparative study, Decis. Support. Syst. 50 (3) (2011) 602–613, https://doi.org/10.1016/j.dss.2010.08.008.
[8] V. Van Vlasselaer, T. Eliassi-Rad, L. Akoglu, M. Snoeck, B. Baesens, GOTCHA! Network-based fraud detection for social security fraud, Manag. Sci. 63 (9) (2017) 3090–3110, https://doi.org/10.1287/mnsc.2016.2489.
[9] P. Castellón González, J.D. Velásquez, Characterization and detection of taxpayers with false invoices using data mining techniques, Expert Syst. Appl. 40 (5) (2013) 1427–1436, https://doi.org/10.1016/j.eswa.2012.08.051.
[10] A. Bonein, La réciprocité, entre psychologie et rationalité économique, Revue Française d'économie 23 (2018) 203–232.
[11] H. Gintis, Strong reciprocity and human sociality, Journal of Theoretical Biology 206 (2000) 169–179, https://doi.org/10.1006/jtbi.2000.2111.
[12] B.S. Frey, B. Torgler, Tax morale and conditional cooperation, J. Comp. Econ. 35 (2007) 136–159, https://doi.org/10.1016/j.jce.2006.10.006.
[13] K. Blaufus, J. Bob, P.E. Otto, N. Wolf, K. Blaufus, J. Bob, P.E. Otto, N. Wolf, T. Effect, 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) (2017) 561–580, https://doi.org/10.1080/09638180.2016.1258319.
[14] R. Calvet, J. Alm, Empathy, sympathy, and tax compliance, J. Econ. Psychol. 40 (2014) 62–82, https://doi.org/10.1016/j.joep.2012.10.001.
[15] J. Lismont, E. Cardinaels, L. Bruynseels, S. De Groote, B. Baesens, W. Lemahieu, J. Vanthienen, Predicting tax avoidance by means of social network analytics, Decis. Support. Syst. 108 (2018) 13–24, https://doi.org/10.1016/j.dss.2018.02.001.
[16] A. Fronzetti Colladon, E. Remondi, Using social network analysis to prevent money laundering, Expert Syst. Appl. 67 (2017) 49–58, https://doi.org/10.1016/j.eswa.2016.09.029.
[17] V. Van Vlasselaer, T. Eliassi-Rad, L. Akoglu, M. Snoeck, B. 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) (2015) 659–666, https://doi.org/10.1145/2808797.2810058.

[18] F. Tian, T. Lan, K.M. Chao, N. Godwin, Q. Zheng, N. Shah, F. Zhang, Mining suspicious tax evasion groups in big data, IEEE Trans. Knowl. Data Eng. 28 (10) (2016) 2651–2664, https://doi.org/10.1109/TKDE.2016.2571686.

[19] M. Barone, M. Coscia, Birds of a feather scam together : trustworthiness homophily in a business network, Soc. Networks 54 (2018) 228–237, https://doi.org/10.1016/j.socnet.2018.01.009.

[20] P. Mehta, J. Mathews, S. Kumar, K. Suryamukhi, C. Sobhan Babu, 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, 2019, pp. 95–109, https://doi.org/10.1007/978-3-030-23551-2_7.

[21] A. Clauset, C. Rohilla Shalizi, M.E.J. Newman, Power-law distributions in empirical data, SIAM Review 51 (4) (2009) 661–703, arXiv:arXiv:0706.1062v2, https://doi.org/10.1214/13-AOAS710.

[22] M. Newman, Networks: An Introduction, Oxford University Press, 2010. arXiv:1212.2425, doi:https://doi.org/10.1093/acprof:oso/9780199206650.001.0001 https://books.google.es/books?hl=es{&}lr={&}id=-DgTDAAAQBAJ{&}oi=fnd{&}pg=PR5{&}dq=complex+networks+newman+2010{&}ots=PA2XnmpPFO{&}sig=920B0s2uWt{_}pXpgdqtzJMXW_CYM.

[23] L.C. Freeman, Centrality in social networks conceptual clarification, Soc. Networks 1 (3) (1978) 215–239, https://doi.org/10.1016/0378-8733(78)90021-7.

[24] G. Csárdi, T. Nepusz, The igraph software package for complex network research, InterJournal Complex Systems 1695 (2006) 1695. http://www.necsi.edu/events/iccs6/papers/c1602a3c126ba822d0bc4293371c.pdf.

[25] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[26] R. Couronné, P. Probst, A.L. Boulesteix, Random forest versus logistic regression: a large-scale benchmark experiment, BMC Bioinformatics 19 (2018) 270, https://doi.org/10.1186/s12859-018-2264-5.

[27] J.J. Levy, J.J. Levy, J.J. Levy, A.J. O'Malley, A.J. O'Malley, Don't dismiss logistic regression: the case for sensible extraction of interactions in the era of machine learning, BMC Med. Res. Methodol. 20 (2020) 1–15, https://doi.org/10.1186/s12874-020-01046-3.

[28] P. Bonacich, P. Lloyd, Eigenvector-like measures of centrality for asymmetric relations, Soc. Networks 23 (3) (2001) 191–201, https://doi.org/10.1016/S0378-8733(01)00038-7.

[29] R.S. Burt, Structural holes and good ideas, Am. J. Sociol. 110 (2) (2004) 349–399, https://doi.org/10.1086/421787.

[30] M.G. Everett, S.P. Borgatti, Unpacking Burt's constraint measure, Soc. Networks 62 (February) (2020) 50–57, https://doi.org/10.1016/j.socnet.2020.02.001.

[31] U. Brandes, A faster algorithm for betweenness centrality, J. Math. Sociol. 25 (2) (2001) 163–177, https://doi.org/10.1080/0022250X.2001.9990249.

[32] C.X. Ling, J. Huang, H. Zhang, AUC: a statistically consistent and more discriminating measure than accuracy, in: IJCAI International Joint Conference on Artificial Intelligence, 2003, pp. 519–524.

[33] A. Fisher, 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.

[34] P. Biecek, DALEX: explainers for complex predictive models in R, The Journal of Machine Learning Research 19(1) (2018) 3245–3249, https://doi.org/10.1016/j.pss.2006.11.016.

[35] J. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (5) (2001) 1189–1232. https://www.jstor.org/stable/2699986.

[36] B.M. Greenwell, pdp: An R package for constructing partial dependence plots, The R Journal 9 (1) (2017) 421–436. https://github.com/bgreenwell/pdp/issues.

Christian González-Martel is Assistant Professor at the Department of Quantitative Methods in Economics and Management at the University of Las Palmas de Gran Canaria. He obtained his Ph.D. in Economics from the University of Las Palmas de Gran Canaria. Christian was a visiting researcher in the Centro Nacional de Supercomputación - Barcelona Supercomputing Center in Barcelona. His main research interest includes machine learning applied in fields like finance, tourism economics, and health economics.

Juan M. Hernández is Full Professor at the Department of Quantitative Methods in Economics and Management fulld at the University of Las Palmas de Gran Canaria, Spain, and the University Institute of Tourism and Sustainable Development (TIDES). Currently, his main line of research is the use of the complex networks perspective for the analysis of real systems (business, tourism, fishery, etc), from both theoretical and empirical points of view.

Casiano A. Manrique-de-Lara-Peñate Assistant Professor at the Department of Applied Economics at the University of las Palmas de Gran Canaria, Spain and member of the University Institute of Tourism and Sustainable Development (TIDES). His main line of research is economic modelling (IO and CGE) and has been involved in research projects related to the general equilibrium and microsimulation modelling of VAT.