Detection Model of Legally Registered Mafia Firms in Italy

Authors: Ravenda, Diego; Argilés-Bsoch, Josep M[a]; Valencia-Silva, Maika M.

Abstract

This paper develops a model that can contribute to the detection of legally registered firms defined as Mafia firms (LMFs) due to having been confiscated by judicial authorities, in relation to alleged connections of their owners with Italian organized crime. The model correctly classifies 76.41% of firms within a matched sample of 852 firm-years including LMFs and lawful firms.

Furthermore, we present an analysis of financial statement characteristics of singular private firms which are socially irresponsible by nature and whose incentives, modus operandi and legal financial statement formats differ from those of listed companies. In particular, we show that specific accruals and earnings management proxies may provide more insight into accounting manipulation patterns of LMFs.

More importantly, our paper can help practitioners and regulators identify accounting signals that can be used in risk assessment models or in the detection of criminal infiltrations and related illicit practices.

Keywords: corporate social responsibility; corruption; earnings management; fraud; legally registered Mafia firms; real activities manipulation.

1 Introduction

The Mafias, which are considered to be the most sophisticated form of criminal organization, also run businesses in the lawful economic sphere in which they usually invest proceeds from illicit trafficking (money laundering). Legally registered Mafia firms (LMFs), according to criminologists' terminology, can be defined as firms that are legally registered and apparently engage in lawful activities but are owned by a Mafia family (Champeyrache, 2004). LMFs differ from lawful firms (LWFs) in three main ways (Gambetta, 1993; Fantò, 1999): the owners are members of a criminal organization; funding partially or totally comes from illegal activities; and criminal methods involving violence, intimidation or corruption might be used while doing business. Legal and illegal activities are therefore closely intertwined within LMFs as the legal activities mostly serve to launder profits stemming from illegal ones (Fantò, 1999).

In this study we examine a sample of 198 Italian legally registered firms defined as LMFs due to having been confiscated at some point by judicial authorities, in relation to alleged connections of their owners with Italian organized crime. In particular, we first analyze whether accounting information of LMFs embeds some significant differences from that of similar firms for which there is no evidence of criminal connection (LWFs). Based on the identified differences, we develop a logistic regression model that can contribute to detecting LMFs and find practical application in forensic accounting. Among the different financial variables we test to predict criminal connections, we particularly focus on earnings management (EM) proxies. Indeed, the large amount of research on EM carried out thus far indicates that managers discretionally manage earnings for different purposes using a wide variety of methods. Specifically, they carry out special transactions, so-called real activities manipulation (RM), that usually affect firm's operating activities, expenses and cash flows from operations (CFO) (e.g., Roychowdhury, 2006) and manipulate discretionary accruals (accrual-based EM) with no CFO impact (e.g., Dechow and Dichev, 2002; Dechow *et al.*, 2010). Hence, in this study we examine both methods of EM as well as developing some new proxies for EM in order to reflect specific characteristics of LMFs.

As far as we know there are no previous studies in the literature that seek to develop an accounting detection model of LMFs. Nonetheless, considering the supposed fraudulent purposes of LMFs such as money laundering and tax evasion we refer to previous studies that develop prediction models of financial statement frauds and related manipulations using financial and non-financial variables (Beneish, 1997; Summers and Sweeney, 1998; Lee *et al.*, 1999; Marquardt and Wiedman, 2004; Erickson *et al.*, 2005; Jones *et al.*, 2008; Brazel *et al.*, 2009; Dechow *et al.*, 2011; Perols and Lougee, 2011). A main difference between LMFs and firms committing financial statement fraud examined in previous research is that in the former the fraudulent purpose is genetic and strictly related to their existence, whereas in the latter fraud is subsequently committed due to specific circumstances.

Overall, our results reveal that our detection model is able to correctly classify 76.41% of firms within a matched sample of 852 firm-year observations including LMFs and LWFs. More specifically, our model detects 76.29% of LMFs (sensitivity) and 76.53% of LWFs (specificity). Out-of-sample tests confirm the robustness of the predictions and an additional analysis shows that undetected LMFs are significantly larger than detected LMFs. Additionally, consistent with previous studies (Beneish, 1997; Dechow *et al.*, 2011) on fraud prediction, our model shows that unadjusted specific accruals have more predictive power than discretionary accruals and a specific RM proxy such as abnormal material expenses is also a significant predictor of criminal connections.

Our study contributes to the accounting literature given that, to our knowledge, it is the first to develop an accounting detection model of LMFs. More importantly, our paper can aid practitioners and regulators in identifying accounting signals that can be used in risk assessment models or in the detection of criminal infiltrations and related illicit practices. Furthermore, our study shows that analysis of specific accruals and RM proxies may provide more insight into EM patterns of LMFs. Finally, it contributes to research on corporate social responsibility (CSR) (e.g., Jenkins, 2006; Guthrie and Durand, 2008; Carroll and Shabana, 2010), indicating that socially irresponsible firms, such as LMFs, tend to engage more in EM.

The remainder of the paper proceeds as follows: section 2 introduces LMFs; section 3 reviews related research; section 4 describes the research design and sample data; section 5 presents empirical results and their discussion; section 6 includes concluding remarks.

2 Legally Registered Mafia Firms

For the purpose of this study, we define "organized crime" according to the legal provision of the article 416-bis of the Italian criminal code. Specifically, art. 416-bis states that:

"A mafia-type association consists of three or more individuals and those who belong to it make use of the power of intimidation afforded by the associative bond and the state of subjugation and criminal silence which derives from it to commit crimes, to acquire directly or indirectly the management or control of economic activities, concessions, authorizations or public contracts and services, either to gain unjust profits or advantages for themselves or for others, or to prevent or obstruct the free exercise of the vote, or to procure votes for themselves or to others at a time or electoral consultation".

Criminal organizations take on new businesses in order to invest and launder significant financial resources coming from illegal activities. In this way, criminal organizations achieve high profits and social consensus by ensuring employment and income for the population in the areas where they exercise control of the territory. Fantò (1999) suggests that the main trait of LMFs is not the type of business run but the nature of the capital accumulation process that led to their formation as well as the strength of intimidation on which they are hinged. This mafiastyle intimidation is a source of surplus value and competitive advantages of LMFs over LWFs. Arlacchi (1983) identifies the following competitive advantages of the LMFs over the LWFs: discouragement of competition (securing goods and raw materials at favorable prices, as well as orders, contracts and commercial outlets using criminal intimidation); wage compression (evasion of social security contributions and insurance, non-payment of overtime, denial of trade union rights); availability of financial resources (investment of huge proceeds coming from illegal activities without bearing the cost of credit).

After the first instance of court confiscation LMFs are entrusted to one or more legal administrators. The legal administration is an institution designed to reinstate the legality, protect and manage confiscated LMFs and avoid their progressive impoverishment. The body currently in charge of the administration and assignment of LMFs definitively confiscated is the Italian agency Agenzia Nazionale Beni Sequestrati e Confiscati (ANBSC).

3 Related Research

3.1 Financial Statement Fraud

In order to identify the most predictive variables of our detection model of LMFs, we mainly refer to previous studies which develop prediction models of financial statement frauds and related manipulations using financial and non-financial variables.

In this regard, Beneish (1997, 1999) estimates a model for detecting earnings manipulation violating generally accepted accounting principles (GAAP) using financial statement variables. He finds a positive relation between aggregate accruals and likelihood of fraud, confirming Dechow *et al.*'s (1996) previous finding. Beneish (1999) considers that a limitation of the model is that it is estimated using financial information for publicly traded companies and cannot be reliably used to study privately-held firms. Lee *et al.* (1999) subsequently find that the excess of earnings over CFO (income increasing accruals) is significantly greater for a sample of 56 firms committing financial statement fraud relative to a broad control sample of non-fraud firms. Marquardt and Wiedman (2004) document that the specific accruals used in EM violating GAAP vary with the context and related incentives, and consequently provide support for the usefulness of examining individual accruals as well as aggregate accruals in specific EM contexts. Jones *et al*. (2008) find that some measures of discretionary accruals have predictive power for fraudulent restatements of financial statements in 118 firms charged by the Securities and Exchange Commission (SEC) between 1988 and 2001. Brazel *et al*. (2009) provide evidence that inconsistencies between nonfinancial measures and financial measure can help detect firms with high fraud risk. More recently, Perols and Lougee (2011), using a sample of 54 fraud and 54 non-fraud firms, show that fraud firms are more likely to have managed earnings in prior years through discretionary accruals. Finally, Dechow *et al*. (2011) analyze the characteristics of firms investigated by the SEC for misstating earnings on various dimensions and find that, at the time of misstatements, accrual quality is low, both financial and nonfinancial measures of performance are deteriorating and financing activities and related off-balance-sheet activities are much more likely.

3.2 Earnings Management within LMFs

In most of the aforementioned studies EM, measured by several proxies, is a significant variable of the prediction model of financial statement frauds. Hence, we expect EM pattern, including both accrual-based EM and RM, to be significantly different between LMFs and LWFs. In particular, we examine both types of EM activities because recent studies suggest that firms choose between the two mechanisms using the technique that is less costly to them (Roychowdhury, 2006; Cohen *et al*., 2008; Cohen and Zarowin, 2010; Badertscher, 2011; Zang, 2012). In this regard, RM, as a departure from optimal operational decisions, is unlikely to increase firms' long-term value. Hence, some managers might find RM particularly costly because their firms face intense competition in the industry (Zang, 2012). However, these considerations may not be applicable to LMFs which usually face a weak market competition and benefit from significant competitive advantages (Arlacchi, 1983; Fantò, 1999). Additional reasons may lead LMFs to engage in EM practices more than LWFs do. First, RM can be used to permanently reduce taxable income, even more effectively than accrual-based EM, by fraudulently removing certain cash flows from the balance sheets. Second, money laundering may require recording fictitious transactions that may lead to EM pattern detected in our proxies. Third, the great availability of financial resources stemming from illegal activities may reduce the need of bank financing and the related incentive to avoid EM practices in order to exhibit an acceptable earnings quality. Finally, a more intensive EM in LMFs may be fostered by the low level of scrutiny from outsiders of these firms compared to LWFs, in connection with the protection ensured by their criminal ties and infiltrators in all spheres of political and institutional life of the country. In this aspect, some analogy might be found with the case of politically connected firms studied by Chaney *et al*. (2011) which engage more in EM than firms lacking such connections. Additionally, previous studies find that a low external monitoring intensity is associated with a higher level of EM (Duellman *et al*., 2013; Wongsunwai, 2013).

3.3 Earnings Management and Corporate Social Responsibility

A further indication on the different EM pattern between LMFs and LWFs may come from some previous research on the relation between CSR and EM. Indeed, LMFs can be assumed to be socially irresponsible based on the widely accepted Carroll's (1979) definition of CSR implying that, in order to meet social expectations, CSR firms work to make a profit, obey the law, behave ethically, and be a good corporate citizen by financially supporting worthy social causes (Carroll, 1991).

In practice, previous studies use a variety of methods to measure CSR. Some of these methods are: reputation indices or databases such as The Kinder, Lydenberg, and Domini (KLD) database (Waddock, 2003; Hong and Andersen, 2011; Kim *et al.*, 2012) which rates US listed companies based on several social dimensions; corporate crime (Davidson and Worrell, 1990; Baucus and Baucus, 1997) and tax avoidance (Lanis and Richardson, 2012; Dowling, 2013) indicators; content analysis of corporate publications on practices regarding environmental, community, employee, and consumer issues (Gray *et al.*, 1995; Turker, 2009); scales measuring the CSR perceptions and values of managers (Singhapakdi *et al.*, 1996; Ruf et *al.*, 1998; Quazi and O'Brien, 2000); scales considering the extent to which businesses meet the economic, legal, ethical, and discretionary responsibilities imposed on them by their stakeholders (Maignan and Ferrell, 2000; Turker, 2009).

We do not directly measure CSR in LMFs. Nonetheless, the CSR measures applied in previous research support our assumption on the social irresponsibility of LMFs. It is noteworthy that previous studies provide inconsistent evidence with mixed implications on the relation between CSR and EM (Chih *et al*., 2008; Prior *et al*., 2008; Gargouri *et al*., 2010; Hong and Andersen, 2011; Kim *et al.*, 2012; Shafer, 2013). Hence, in this study we aim to provide additional insight into this relation.

4 Research Design

4.1 Variable Definition

Consistent with previous studies on business failure prediction (Dambolena and Khoury, 1980; Karels and Prakash, 1987; Balcaen and Ooghe, 2006; Åstebro and Winter, 2012), we explore a wide range of financial characteristics of LMFs as well as their EM behavior in order to build the best detection model.

4.1.1 Earnings Management Variables

Prior studies (Roychowdhury, 2006; Cohen *et al*., 2008; Cohen and Zarowin, 2010; Badertscher, 2011; Kim *et al*., 2012; Zang, 2012) use different proxies for RM including: abnormal levels of CFO, abnormal productions costs, abnormal discretionary expenses (R&D, advertising, and selling, general, and administrative expenditures). In Italy legal format of income statement classifies expenses by nature rather than by function and production costs cannot be distinguished from discretionary expenditures. Therefore, we adopt three new measures of RM as well as the usual abnormal CFO (*ABCFO*): abnormal material expenses (*ABMAT*), including both raw materials and trading goods, abnormal service expenses (*ABSERV*) and abnormal personnel expenses (*ABPER*).

In LMFs we expect higher *ABMAT* due to fraudulent sales underreporting and the record of fictitious transactions with related parties in order to disguise money laundering and evade taxes. Furthermore, we expect lower *ABPER* due to wage compression practices (Arlacchi, 1983) including evasion of social security contributions. Finally, we expect lower *ABSERV* given that LMFs may be less prone to contract external services (advertising, consultancy, maintenance etc.) because of their aforementioned competitive advantages.

As a measure of accrual-based EM we calculate discretionary aggregate accruals (*DAC*), discretionary revenue accruals (*DREV*) and a new measure of discretionary expense accruals (*DEXP*). Indeed, we consider that LMFs may simultaneously manipulate revenues and expenses and the related cumulative effect may not be detected in aggregate discretionary accrual models which do not provide information as to which components of earnings firms manage and how the EM is achieved (Marquardt and Wiedman, 2004).

Previous studies find that discretionary accrual models have less power to identify manipulation than unadjusted accrual measures supplemented with other financial statement ratios (Beneish, 1997; Dechow *et al*., 2011). Hence, we additionally test in our model, deflated by lagged total assets, unadjusted aggregate accruals (*ACCR*) and some unadjusted specific accruals that are more likely to be manipulated such as: change in receivables (*CH_REC*), change in inventory (*CH_INV*) and change in payables (*CH_PAY*). Following the same reasoning for accrual-based EM we also examine unadjusted proxies for RM by including in our model personnel, material and service expenses deflated by lagged total assets in order to determine whether they show more predictive power than commonly used abnormal RM measures.

4.1.2 Other Variables

Besides accrual-based EM and RM measures we test in our model the following variables, grouped by category, used in prior works on fraudulent financial statements and adapted to the singularities of LMFs.

Asset composition. Previous studies (Loebbecke *et al*., 1989; Persons, 1995; Summer and Sweeney, 1998; Beneish, 1999; Dechow *et al*., 2011) examine asset composition with special regard to receivables and inventories that can be an easy target for manipulation due to the subjective judgment involved in their valuation. Accordingly, we measure asset composition with variables *CATA* (current assets/total assets), *RECTA* (receivables/total assets), *INVTA* (inventory/total assets) and *INTA* (intangible assets/total assets). In comparison with LWFs in the same industry, we expect LMFs to exhibit higher receivables to account for incoming dirty money and lower inventory to avoid taxes (VAT and income tax) through stock underreporting and fictitious purchase transactions.

Performance. We examine some variables expressing the reported firm financial performance and try to detect inconsistencies and signals of possible fraudulent manipulations. In particular, previous fraud research finds that firms that increase revenue fraudulently are more likely to have abnormally high sales growth rates (Erickson *et al*., 2006; Brazel *et al*., 2009). As firms use resources to generate sales, unusual relations between sales and resources used, such as assets (capital productivity) and employees (labor productivity) may be a signal of fraud. Therefore, in line with previous studies (Fanning and Cogger, 1998; Perols and Lougee, 2011), we include *Revenue to Assets* (*REVTA*) and *Revenue to Employee (REVEMPL)* as predictors in our model. We predict a negative relation between *REVTA* and probability of criminal connection (*CRIME*) given that in LMFs revenue may be underreported for tax evasion and there may be a need to quickly overinvest in assets financial resources coming from illicit sources without demanding an immediate competitive return. On the other hand, higher values of *REVEMPL* for LMFs relative to LWFs may be due not only to a fraudulent revenue manipulation but also to the underreporting of the number of employees because of the employment of undeclared workers.

We additionally test Return on Assets (*ROA*) as a predictor given that we expect LMFs to be less profitable than LWFs. Indeed, LMFs may downward manage earnings to avoid tax as well as being oversized and poorly managed. Change in *ROA* (ABS_*CH_ROA*) is also added following Dechow *et al*. (2011) although, differently from the latter, we consider the absolute value in order to reflect higher opportunistic profitability fluctuations not reflecting the actual business performance. In accordance with this higher volatility pattern in LMFs, we furthermore include and expect higher values for absolute changes in percentages of personnel (*ABS_CH_PERSREV*), material (*ABS_CH_MATREV*) and service expenses

(*ABS_CH_SERVREV*) over sales and absolute changes in net income (*ABS_CH_NI*) and CFO (*ABS_CH_CFO)* deflated by lagged total assets.

In line with previous fraud research (Fanning and Cogger, 1998; Summers and Sweeney, 1998; Beneish, 1999; Lee *et al*., 1999) we additionally include the annual absolute change in the ratio receivables to sales (*ABS_CH_RECREV*) also called days' sales in receivables. A significant variation in days' sales in receivables could be the result of a change in credit policy but it may also be suggestive of a fraudulent revenue manipulation (Beneish, 1999). As we expect revenue manipulation to be either upwards or downwards we consider the absolute value of ratio variation. In order to detect a possible simultaneous expense manipulation, we also add a variable for the absolute change in payables to purchases (ABCH*_PAY_EXP).*

Debt. As regards the indebtedness, we expect a positive relation between leverage (*LEV*) (total liabilities/total assets) and *CRIME*. LMFs may be more indebted than LWFs because they may report fictitious business transactions or may obtain favorable payment terms from suppliers using the strength of criminal intimidation (Arlacchi, 1983; Fantò, 1999). More specifically, LMFs may prefer fictitious debt transactions to inject dirty money since regular contributions of capital from shareholders may raise suspicions on their origins. Nonetheless, we expect LMFs to show less bank indebtedness *(LEVBANK*) compared to the rest of LWFs because their access to alternative illegal source of funding may replace bank support.

Liquidity. Regarding liquidity we include current ratio (*CRATIO*: current assets/current liabilities) (Shih *et al*., 2011) and the absolute value of its annual change (*ABS_CH_CRATIO).* We expect a worse and more fluctuating liquidity situation for LMFs given that current assets and liabilities balances may include fictitious fraudulent transactions, undermining the adequacy of these ratios to reflect the actual short-term debt-paying ability of the firms.

Growth. Previous research finds that the fast growth of a firm is an important warning of financial information fraud (Loebbecke *et al.*, 1989; Beasley, 1996; Bell and Carcello, 2000; Shih *et al*., 2011). Consistently, we include percentage increase of total assets (*GROWTH*) as a predictor in our model. Indeed, we expect LMFs to have a higher growth rate than LWFs because of the continuous investment of financial resources stemming from illegal activities.

Non-financial. Following Dechow *et al*. (2011) we add a measure of difference of percentage change in total assets less percentage change in number of employees (*DIF_GROWTH_EMPL*) under the assumption that physical assets and employees are complements and should follow a similar growth pattern. We expect this measure to be significantly lower for LMFs because, although they may overinvest to launder dirty money, a sustained underreporting of number of employees may result in higher fluctuations in the number of employees and higher employee growth rates. Lastly, we include personnel expenses per employee *(PERSEMPL)* expecting a lower value for LMFs due not only to lower remunerations but also to the payment of undeclared envelope wages (Williams, 2009).

4.2 Earnings Management Variable Construction

We need to build measures of accrual-based EM and RM to input as independent variables in our prediction model. Hence, we calculate discretionary accruals (*DAC*) as the residuals from the following Eq. (1) based on the modified Jones model (Dechow *et al*., 1995) with a control for performance (Kothari *et al*., 2005):

$$\frac{ACCR_t}{TA_{t-1}} = \beta_0 + \beta_1 \frac{1}{TA_{t-1}} + \beta_2 \frac{\Delta REV_t - \Delta AR_t}{TA_{t-1}} + \beta_3 \frac{PPE_t}{TA_{t-1}} + \beta_4 ROA_{t-1} + \varepsilon_t \qquad (1)$$

Where in year t (or t - 1), *ACCR* denotes total accruals; *TA*, Δ*REV*, Δ*AR*, *PPE*, and *ROA* represent total assets, changes in net revenue, changes in accounts receivables, property, plant, and equipment, and return on assets, respectively. Parameters of Eq. (1) are estimated cross-sectionally for each industry-year with at least 15 observations in order to control for industry-wide changes under different economic conditions (Jeter and Shivakumar, 1999) that affect total accruals while allowing the coefficients to vary across time (DeFond and Jiambalvo, 1994; Kasznik, 1999). We use all active firms in AIDA (excluding LMFs) which are not listed on the stock exchange and with financial statements available for 10 years from 2003 to 2012.

Consistent with previous studies of EM (Healy, 1985; Jones, 1991; Dechow *et al*., 1995), *ACCR* are computed as:

$$ACCR_t = \Delta CA_t - \Delta CL_t - \Delta CASH_t + \Delta STD_t - DEP_t \tag{2}$$

Where:

ΔCA = change in current assets, ΔCL = change in current liabilities, $\Delta CASH$ = change in cash and cash equivalents, ΔSTD =change in debt included in current liabilities and *DEP* = depreciation and amortization expenses.

CFO is computed as:

$$CFO = Earnings\ before\ tax - ACCR \tag{3}$$

Following Caylor (2010) and Stubben (2010), we calculate discretionary revenue accruals (*DREV)* and a new measure of discretionary expense accruals (*DEXP*) as the residual from the following Eq. (4) estimated in the same way as *DAC.*

$$\frac{\Delta AR_t(\Delta AP_t)}{TA_{t-1}} = \beta_0 + \beta_1 \frac{1}{TA_{t-1}} + \beta_2 \frac{\Delta REV_t}{TA_{t-1}} + \beta_3 \frac{\Delta CFO_{t+1}}{TA_{t-1}} \varepsilon_t \tag{4}$$

Where ΔAP represents change in accounts payables.

Furthermore, we calculate the abnormal level of material expenses (*ABMAT)* and personnel expenses (*ABPER*) as the estimated residual of the following model adopted by Roychowdhury (2006) for production costs:

(5)

$$\frac{MAT_t(PER_t)}{TA_{t-1}} = \beta_0 + \beta_1 \frac{1}{TA_{t-1}} + \beta_2 \frac{S_t}{TA_{t-1}} + \beta_3 \frac{\Delta S_t}{TA_{t-1}} + \beta_4 \frac{\Delta S_{t-1}}{TA_{t-1}} + \varepsilon_t$$

Where MAT_t and PER_t are respectively material expenses and personnel expenses in year t that we assume mostly related to production; S_t is the net sales in year t; and ΔS_t is the change in net sales from year t-1 to t (S_t - S_{t-1}). Eq. (5) is estimated in the same way as *DAC*.

Additionally, we estimate the abnormal level of service expenses (*ABSER*V) as the residual from the following Eq. (6) used by Roychowdhury (2006) for discretionary expenses and estimated in the same way as *DAC*:

$$\frac{SERV_t}{TA_{t-1}} = \beta_0 + \beta_1 \frac{1}{TA_{t-1}} + \beta_2 \frac{S_{t-1}}{TA_{t-1}} + \varepsilon_t \quad (6)$$

Finally, in line with Dechow *et al*. (1998) and Roychowdhury (2006) we estimate abnormal CFO (*ABCFO*) as the residual from the following Eq. (7) estimated in the same way as *DAC*:

$$\frac{CFO_t}{TA_{t-1}} = \beta_0 + \beta_1 \frac{1}{TA_{t-1}} + \beta_2 \frac{S_t}{TA_{t-1}} + \beta_3 \frac{\Delta S_t}{TA_{t-1}} + \varepsilon_t \quad (7)$$

4.3 Detection Model

In order to build our detection model we start with the estimation of the following logistic regression model (Eq. 8) where the dependent dummy variable *CRIME* takes a value of 1 for LMFs and 0 for LWFs:

$$Pr\ (CRIME) = f\ (EM\ variables,\ Asset\ composition\ variables,\ Performance\ variables,\ Debt\ variables,\ Liquidity\ variables,\ Growth\ variable,\ Nonfinancial\ variables) \quad (8)$$

Following a similar approach adopted by Dechow *et al*. (2011) for prediction of accounting misstatements, we group the variables in different categories. Table 1 describes the independent variables and their calculation, classifies them by category and indicates their predicted sign as previously discussed.

(Insert Table 1 approximately here)

4.4 Data and Sample Selection

LMFs sample consists of 198 firms confiscated to organized crime, some of them provided by ANBSC and others found in online newspapers and AIDA database. The financial statements for all firms are obtained from AIDA, the Italian Bureau Van Dijk database. It contains comprehensive information on 1 million companies with a turnover above € 500,000 in Italy, including the indication for some of them of the confiscation status and date of confiscation. Firms provided by ANBSC have all been confiscated by final judgment but their small size or their liquidation means that only 54 out of 1,663 have financial statements available on AIDA. In addition, we include firms confiscated in first instance and found in AIDA database (118) and online newspapers (52) until reaching a total of 224. We only consider firm-year observations prior to the confiscation year as once confiscated and subject to legal administration LMFs may lose their distinctive characteristics. Hence, out of these 224 LMFs we eliminate 26 confiscated before 2005 whose needed financial statements are unavailable on AIDA which only includes years from 2003 to 2012. Finally, we end up with a sample of 198 LMFs. Moreover, some missing data on AIDA for the calculation of several tested variables in some years further reduce the number of firm-year observations in the final detection model which ends up being 426.

Table 2 presents the industry distribution by two-digit SIC groups of LMFs in our sample and AIDA population of active unlisted firms with available financial data from 2003 to 2012 in the same industries as LMFs.

(Insert Table 2 approximately here)

Compared to the population of active and unlisted firms in AIDA with available financial data from 2003 to 2012, the sample LMFs are especially more abundant in industry groups: building construction-general contractors and operative builders (16.67% of criminal sample versus 7.00% of population), food stores (7.58% versus 2.22%) and Motor freight transportation and warehousing (9.09% versus 3.69%). On the other hand, there is a lower proportion of LMFs mostly in wholesale trade, durable goods (11.11% versus 17.95%), business services (1.01% versus 6.38%) and fabricated metal products, except machinery and transportation equipment (0.51 versus 8.98%).

In order to build our full sample for the model estimate, we use a matched sample design (Defond and Jiambalvo, 1994; Perry and Williams, 1994; Defond and Subramanyam, 1998; Teoh *et al*., 1998; Kothari *et al*., 2005). Specifically, we match each LMF-year with a LWF-year on fiscal reporting year, industry and size proxied by total assets.

Table 3 summarizes the sample selection procedure that yields the 198 LMFs and the 418 control LWFs.

(Insert Table 3 approximately here)

Table 4 includes number of LMFs by confiscation year. It can be seen that 2012 is the year with largest number of confiscated firms and more than 50% of firms have been confiscated from 2011 to 2013.

(Insert Table 4 approximately here)

5 Results and Discussions

5.1 Descriptive Statistics and Univariate Analysis

Table 5 presents descriptive statistics for each variable considered for the development of our detection model comparing LMF-years to their matched LWF-years before confiscation. All continuous variables are winsorized at the top and bottom 1 percent of their distributions to avoid the influence of outliers.

(Insert Table 5 approximately here)

As regards accrual-based EM variables, it is noteworthy that, as expected, *ABSDAC*, *ABSDREV* and *ABSDEXP* are significantly (p<0.01) higher for LMFs relative to LWFs, suggesting a higher degree of aggregate accrual-based, revenue-accrual based and expense accrual-based EM, respectively. As regards RM variables, variable *ABMAT* is positive and significantly (p<0.01) higher for LMFs indicating an income-decreasing RM that is offset by an income-increasing RM suggested by significantly (p<0.05) lower variables *ABPER* and *ABSERV*. Significantly (p<0.05) lower variable *ABCFO* for LMFs provides evidence that the cumulative effect of RM is a reduction of *CFO*. As regards unadjusted EM proxies, variables *CH_REC*, *CH_INV* and *CH_PAY*, each representing a different specific accrual, are significantly (p<0.05) higher for LMFs. Similar to the results of related RM proxies, variables *PERTA* and *SERVTA* are significantly (p<0.01) lower and variable *MATTA* is significantly (p<0.01) higher for LMFs. As regards asset composition variables, marginally significantly (p<0.10) higher variable *CATA* documents a higher liquidity in asset composition of LMFs. This is partially due to higher receivables, as showed by significantly (p<0.05) higher variable *RECTA*, and despite the significantly (p<0.05) lower variable *INVTA*. As far as performance variables are concerned, profitability variable *ROA* is significantly (p<0.01) lower for LMFs and more volatile, as suggested by significantly (p<0.05) higher variable *ABS_CH_ROA*. Significantly (p<0.05) lower variables *SERVREV* and *PERSREV* and significantly (p<0.01) higher variable *MATREV* for LMFs provide further evidence on lower service and personnel expenses and higher material expenses with respect to sales, respectively. Significantly (p<0.01) higher variables *ABS_CH_PERSREV*, *ABS_CH_MATREV*, *ABS_CH_SERVREV*, *ABS_CH_NI*, *ABS_CH_RECREV* and *ABS_CH_PAYEXP* for LMFs provide further evidence on the higher volatility of their reported performance which foster suspicions on opportunistic and fraudulent manipulations. As expected, LMFs are significantly (p<0.01) more leveraged (*LEV*), although their bank indebtedness (*LEVBANK*) is significantly (p<0.01) lower. Variable

CRATIO is significantly (p<0.01) lower for LMFs indicating a theoretical weakness in the ability to meet their short term debt obligations. It is worth noting the expected significantly (p<0.01) higher total assets growth rate (*GROWTH*) of LMFs. Finally, according to our expectations, non-financial variable *PERSEMPL* is significantly (p<0.01) lower for LMFs providing indication of wage compression practices (Arlacchi, 1983; Fantò, 1999) and non-financial variable *DIF_GROWTH_EMPL* is negative and significantly (p<0.01) lower for LMFs.

Table 6 displays Pearson correlations among EM related variables taken into account for developing our detection model. High correlations identified among some variables warn against their simultaneous inclusion in the detection model.

(Insert Table 6 approximately here)

5.2 Multivariate Logistic Regression Analysis

We estimate a cross-sectional logistic regression to determine whether the variables examined in univariate tests are jointly significant in detecting LMF-years. We use a stepwise backward elimination technique to arrive at a parsimonious model that best predicts LMFs within our sample. The model is displayed in Table 7.

(Insert Table 7 approximately here)

The chi-square test indicates the significance of the overall model. As showed at the bottom of the Table 7, using a probability cut-off point of 0.50 the model correctly classifies 76.29% of the total LMF-years (sensitivity) and 76.53% of the total LWF-years (specificity) with a total rate of 76.41 firm-years correctly classified. Similar to previous studies (Lisowsky, 2010; Dimmock and Gerken, 2012; Åstebro and Winter, 2012), to illustrate the possible tradeoffs between false positives and correctly predicted LMFs at various probability cutoff-points, Fig. 1 shows a receiver operating characteristic (ROC) curve for the detection model. The area under the ROC Curve of our estimated model is approximately 0.82, indicating strong discriminatory power of the model to identify LMFs (Hosmer and Lemeshow, 2000).

(Insert Figure 1 approximately here)

Fig. 2 shows the graph of sensitivity and specificity for each probability cut-off point for the detection model of Table 7. A reduction of the cut-off from 0.50 to 0.35 scores a sensitivity of approximately 90% and a specificity of approximately 60%. Indeed, considering the higher misclassification cost for LMFs relative to LWFs, reducing the cut-off point from 0.5 might be a convenient option.

(Insert Figure 2 approximately here)

Turning to the results of the estimated detection model in Table 7, it is noteworthy that, within the accrual-based EM variables, coefficients on *CH_REC* and *CH_INV* are negative and significant ($p<0.05$) supporting previous studies which find that unadjusted specific accruals have more power to identify fraudulent manipulations than discretionary accruals (Beneish, 1997; Dechow *et al*., 2011). The former are thus preferable because of the fewer calculation efforts they require. Regarding variable *ABMAT,* its coefficient is positive and significant ($p<0.01$), as expected, providing evidence that LMFs are more likely to upward manage material expenses than LWFs do. On the other hand, positive and significant ($p<0.01$) coefficient on *INTA* and negative and significant ($p<0.05$) coefficient on *INVTA* respectively suggest that LMFs are more likely to report higher intangible assets and lower inventory with respect to total assets.

As far as performance variables are concerned, coefficient on *REVTA* is negative and significant ($p<0.01$) as expected. Furthermore, negative and significant ($p<0.01$) coefficient on *SERVREV* suggests lower service expenses with respect to sales in LMFs. For the rest of variables of the models, the results of univariate tests are mostly confirmed and the same considerations apply. Some exceptions are variables *ABS_CH_MATREV* and *ABS_CH_NI* whose coefficients are not significant at conventional levels in spite of improving the predictive power of the model. Another exception is the variable *REVEMPL* whose coefficient is positive and significant ($p<0.05$) apparently suggesting a higher labor productivity in LMFs relative to LWFs. Nonetheless, we are more inclined to believe that this result is mainly due to the underreporting of number of employees.

Finally, Table 8 shows the illicit activities which may be reflected by the variables included in the final detection model. Our analysis is mostly based on the assumptions made in the variable definition section 4.1. Money laundering as well as labor, income and value added tax evasion are assumed to be the primary incentives which should be considered whether additional variables are included in the model in order to improve its predictive power.

(Insert Table 8 approximately here)

5.3 Robustness Tests

In this subsection we test whether the within-sample predictions are robust out-of-sample through a cross-validation. For this purpose we estimate three detection models excluding in turns LMFs confiscated in each year between 2011 and 2013 with their control firms and predicting values for each excluded hold-out sample. Related estimates and detection accuracy rates for each yearly hold-out sample are presented in Table 9.

(Insert Table 9 approximately here)

The results indicate that the overall predictive power of the models at cut-off of 0.50 is 69.55%, 71.57% and 77.91% in the hold-out samples of LMFs confiscated in 2013, 2012 and 2011, respectively. Due to the relatively small difference from our tested model we consider that the out-of-sample tests support the robustness of our detection model.

5.4 Analysis of Undetected LMFs

We perform a further analysis of LMFs undetected by our model in order to determine whether they present some significant differences from detected LMFs.

Table 10 shows the industry distribution of undetected and detected LMFs. An untabulated Pearson Chi-squared test of independence indicates that industry distribution of undetected LMFs is not significantly different from that of detected LMFs.

(Insert Table 10 approximately here)

Table 11 presents univariate tests of differences between undetected and detected LMFs including detection model variables and two additional variables measuring firm size.

(Insert Table 11 approximately here)

It is noteworthy that undetected LMFs are significantly ($p<0.05$) larger than detected LMFs in terms of both logarithm of total assets and number of employees. Indeed, larger firms are more easily scrutinized by regulators (Siregar and Utama, 2008) and may have more resources and incentives to better disguise illicit practices by enhancing the rationality and economic credibility of accounting information (Compin, 2008). Interestingly, as regards detection model variables, *ABMAT* is significantly ($p<0.01$) lower for undetected LMFs. Furthermore, undetected LMFs exhibit a significantly ($p<0.01$) lower total indebtedness (*LEV*) and a significantly ($p<0.01$) higher bank indebtedness (*LEVBANK*). Finally, significantly ($p<0.01$) higher variables *PERSEMPL* and *DIF_GROWTH_EMPL* for undetected LMFs may indicate less adoption of wage compression practices (Arlacchi, 1983).

6 Conclusions

In this study we develop a logistic regression model that can contribute to the detection of LMFs in Italy based on their financial statement characteristics. Overall, our results reveal that our model is able to detect 76.29% of LMF-years (sensitivity) and 76.53% of LWF-years (specificity) within a matched sample of 852 firm-years including both LMFs and LWFs.

As a primary contribution, our paper can aid practitioners and regulators in identifying accounting signals that can be used in risk assessment models or in the detection of criminal infiltrations and related illicit practices. In particular, a high probability score resulting from the model could be used as a further selection criterion of firms to be regularly inspected in order to unmask illegal activities and as a red flag strengthening existing evidence of Mafia activities. Indeed, because of its limitations, the model cannot by itself support allegations of Mafia infiltrations within a firm without additional proofs.

We recognize that in the future our detection model might need to be adapted to the continuous evolution of Mafia practices. Nonetheless, we do not expect any significant change in the practices of LMFs as an immediate reaction aiming to undermine the effectiveness of an auditing procedure based on our model. Indeed, LMFs are already engaged in disguising their illicit practices and the patterns disclosed by our model are a necessary consequence of these attempts. Furthermore, confiscations of LMFs are mostly based on investigations carried out by authorities on parallel illicit activities and criminal bonds of the owners that significantly benefit LMFs by granting them sources of funding and business opportunities. The imputation of the owners for mafia-type association automatically implies the confiscation of all their assets including firms. Hence, a change in the internal LMFs practices would not prevent authorities from accomplishing their investigations.

However, our findings are subject to several limitations. We cannot be completely sure that control sample LWFs are not connected to criminal organizations despite having never been confiscated. Nonetheless, considering the large population of 78,340 firms from which control sample LWFs have been selected, we assume a very low probability of a significant presence of LMFs in our control sample. Although we conduct extensive out-of-sample tests, we cannot reject the possibility that our detection model is biased because undetected LMFs are unobservable and smaller LMFs unavailable on AIDA are excluded. Furthermore, there could be selection biases in LMFs pursued by Italian authorities.

We propose several opportunities for future research. First, other detection techniques (multiple discriminant analysis, neural networks, decision trees, etc.) could be tested in order to find out whether they perform better than our logistic model. Second, additional financial and non-financial information from other sources may be considered to improve the predictive power of the model. Third, the model could be applied to other types of illegal firms such as simple tax evaders that, although not directly connected to any criminal organization, may have behavior patterns similar to LMFs. Finally, this study could be replicated in other countries, where organized crime is deeply rooted, in order to determine whether the results are confirmed in a different cultural, legal and institutional context.

7 References

Arlacchi, P., 1984, *La mafia imprenditrice: L'etica mafiosa e lo spirito del capitalismo.* Bologna: Il Mulino.

Åstebro, T. and J. K. Winter, 2012, "More than a Dummy: The Probability of Failure, Survival and Acquisition of Firms in Financial Distress". *European Management Review,* **9**: 1-17.

Badertscher, B. A., 2011, "Overvaluation and the choice of alternative earnings management mechanisms". *The Accounting Review,* **86**: 1491–1518.

Balcaen, S. and H. Ooghe, 2006, "35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems". *The British Accounting Review,* **38**: 63–93.

Baucus, M. S. and D. A. Baucus, 1997, “Paying the Piper: An Empirical Examination of Longer-Term Financial Consequences of Illegal Corporate Behavior”. *Academy of Management Journal*, **40**: 129–151.

Bhattacharya, U., H. Daouk and M. Welker, 2003, “The World Price of Earnings Opacity”. *The Accounting Review*, **78**: 641–678.

Beasley, M. S., 1996, “An empirical analysis of the relation between the board of director composition and financial statement fraud”. *Accounting Review*, **71:** 443-465.

Bell, T. B. and J. V. Carcello, 2000, “A decision aid for assessing the likelihood of fraudulent financial reporting”. *Auditing*, **19(1)**: 168-184.

Beneish, M. D., 1997, “Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance”. *Journal of Accounting and Public Policy,* **16**: 271–309.

Beneish, M. D., 1999, “The detection of earnings manipulation”. *Financial Analysts Journal,* **55(5)**: 24–36.

Brazel, J. F., K. L. Jones and M. F. Zimbelman, 2009, “Using nonfinancial measures to assess fraud risk”. *Journal of Accounting Research*, **47**: 1135−1166.

Carroll, A. B., 1979, “A three-dimensional conceptual model of corporate performance”. *The Academy of Management Review*, **4**: 497–505.

Carroll, A. B., 1991, “The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders”. *Business Horizons,* **34**: 39–48.

Carroll, A. B. and K. M. Shabana, 2010, "The Business Case for Corporate Social Responsibility: A Review of Concepts, Research and Practice". *International Journal of Management Reviews,* **12**: 85-105.

Caylor, M. L., 2010, “Strategic revenue recognition to achieve earnings benchmarks”. *Journal of Accounting and Public Policy,* **29**: 82–95.

Champeyrache C., 2004, *Entreprise Légale, Propriétaire Mafieux: Comment la Mafia Infiltre l'Economie Légale*. 1st. edn. Paris: Editions CNRS.

Chaney, P. K., M. Faccio and D. Parsley, 2011, "The quality of accounting information in politically connected firms". *Journal of Accounting and Economics*, **51:** 58-76.

Chih, H., C. Shen and F. Kang, 2008, "Corporate social responsibility, investor protection, and earnings management: Some international evidence". *Journal of Business Ethics*, **79**: 179–198.

Cohen, D. A., A. Dey and T. Z. Lys, 2008, "Real and accrual-based earnings management in the pre- and post-Sarbanes-Oxley periods". *The Accounting Review*, **83**: 757–787.

Cohen, D. A. and P. Zarowin, 2010, "Accrual-based and real earnings management activities around seasoned equity offerings". *Journal of Accounting and Economics*, **50**: 2–19.

Compin, F., 2008, "The role of accounting in money laundering and money dirtying". *Critical Perspectives on Accounting*, **19**: 591–602.

Dambolena, I., S. Khoury, 1980, "Ratio stability and corporate failure". *Journal of Finance,* **33**: 1017–1026.

Davidson, W. N. and D. L. Worrell, 1990, "A Comparison and Test of the Use of Accounting and Stock Market Data in Relating Corporate Social Responsibility and Financial Performance". *Akron Business and Economic Review*, **21**: 7–19.

Dechow, P. M. and I. D. Dichev, 2002, "The quality of accruals and earnings: The role of accrual estimation errors". *The Accounting Review,* **77**: 35-59.

Dechow, P. M., W. Ge, C. R. Larson and R. G. Sloan, 2011, "Predicting material accounting misstatements". *Contemporary Accounting Research*, **28**: 17-82.

Dechow, P. M., S. P. Kothari and R. L. Watts, 1998, "The relation between earnings and cash flows". *Journal of Accounting and Economics,* **25**: 133–168.

Dechow, P. M. and D. J. Skinner, 2000, “Earnings management: Reconciling the views of accounting academics, practitioners, and regulators”. *Accounting Horizons*, **14**: 235–50.

Dechow, P. M., R. G. Sloan and A. P. Sweeney, 1995, “Detecting Earnings Management”. *The Accounting Review*, **70**: 193–226.

Dechow, P. M., R. G. Sloan and A. P. Sweeney, 1996, “Causes and consequences of earnings manipulation: an analysis of firms subject to enforcement actions by the SEC”. *Contemporary Accounting Research*, **13**: 1–36.

DeFond, M. L. and J. Jiambalvo, 1994, “Debt covenant violation and manipulation of accruals”. *Journal of Accounting and Economics*, **17**: 145–176.

DeFond, M. L. and K. R. Subramanyam, 1998, “Auditor changes and discretionary accruals”. *Journal of Accounting and Economics,* **25**: 35–67.

Dimmock, S. G. and W. C. Gerken, 2012, “Predicting fraud by investment managers”. *Journal of Financial Economics*, **105**: 153-173.

Dowling, G. R., 2013, “The curious case of corporate tax avoidance: Is it socially irresponsible?” *Journal of Business Ethics*, 1-12.

Duellman, S., A. S. Ahmed and A. M. Abdel-Meguid, 2013, “An empirical analysis of the effects of monitoring intensity on the relation between equity incentives and earnings management”. *Journal of Accounting and Public Policy*, **32**: 495-517.

Dyreng, S. D., M. Hanlon and E. L. Maydew, 2012, “Where do firms manage earnings?” *Review of Accounting Studies,* **17**: 649–687.

Erickson, M., M. Hanlon and E. L. Maydew, 2006, “Is there a link between executive equity incentives and accounting fraud?” *Journal of Accounting Research*, **44**: 113−143.

Fanning, K. and K. Cogger, 1998, “Neural network detection of management fraud using published financial data”. *International Journal of Intelligent Systems in Accounting, Finance and Management*, **7**: 21−41.

Fantò E., 1999, *L'impresa a partecipazione mafiosa. Economia Legale ed Economia Criminale.* 1st. edn. Bari: Dedalo.

Gambetta D., 1993, *The Sicilian Mafia: the business of private protection.* 1st. edn. Cambridge: Harvard University Press.

Gargouri, R., C. Francoeur and R. Shabou, 2010, "The relation between corporate social performance and earnings management". *Canadian Journal of Administrative Sciences*, **27**: 320–334.

Gray, R., R. Kouhy and S. Lavers, 1995, "Corporate social and environmental reporting: a review of the literature and a longitudinal study of UK disclosure". *Accounting, Auditing & Accountability Journal,* **8**: 47-77.

Guthrie, D. and D. Durand, 2008, "Social issues in the study of management". *European Management Review*, **5**: 137-149.

Healy, P. M., 1985, "The effect of bonus schemes on accounting decisions". *Journal of Accounting and Economics*, **7**: 85–107.

Healy, P. M. and J. M. Wahlen, 1999, "A review of the earnings management literature and its implications for standard setting". *Accounting Horizons*, **13**: 365−383.

Hong, Y. and M. L. Andersen, 2011, "The relationship between corporate social responsibility and earnings management: An exploratory study". *Journal of Business Ethics*, **104**: 461-471.

Hosmer, D. and S. Lemeshow, 2000, *Applied Logistic Regression*. 2nd edition. New York, NY: John Wiley & Sons, Inc.

Jenkins, H., 2006, "Small Business Champions for Corporate Social Responsibility". *Journal of Business Ethics*, **67**: 241-256.

Jeter, D. C. and L. Shivakumar, 1999, "Cross-sectional estimation of abnormal accruals using quarterly and annual data: Effectiveness in detecting event-specific earnings management". *Accounting and Business Research*, **29:** 299-319.

Jiambalvo, J., 1996, "Discussion of "Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC"". *Contemporary Accounting Research*, **13**: 37-47.

Jones, J., 1991, "Earnings management during import relief investigations". *Journal of Accounting Research*, **29**: 193–228.

Jones, K. L., G. V. Krishnan and K. D. Melendrez, 2008, "Do models of discretionary accruals detect actual cases of fraudulent and restated earnings? An empirical analysis". *Contemporary Accounting Research*, **25**: 499-531.

Karels, G. V. and A. J. Prakash, 1987, "Multivariate normality and forecasting of business bankruptcy". *Journal of Business Finance & Accounting*, **14**: 573–593.

Kasznik, R., 1999, "On the association between voluntary disclosure and earnings management". *Journal of Accounting Research*, **33**: 57-81.

Kim, Y., M. S. Park and B. Wier, 2012, "Is Earnings Quality Associated with Corporate Social Responsibility?" *The Accounting Review*, **87**: 761–796.

Kothari, S. P., A. Leone and C. Wasley, 2005, "Performance matched discretionary accrual measures". *Journal of Accounting and Economics*, **39**: 163–197.

Lanis, R. and G. Richardson, 2012, "Corporate social responsibility and tax aggressiveness: An empirical analysis". *Journal of Accounting and Public Policy*, **31**: 86–108.

Lee, T. A., R. W. Ingram and T. P. Howard, 1999, "The difference between earnings and operating cash flow as an indicator of financial reporting fraud". *Contemporary Accounting Research*, **16**: 749-786.

Lisowsky, P., 2010, "Seeking shelter: Empirically modeling tax shelters using financial statement information". *Accounting Review*, **85**: 1693-1720.

Loebbecke, J. K., M. M. Eining and J. J. Willingham, 1989, “Auditors' Experience with Material Irregularities: Frequency, Nature, and Detect Ability”. *Auditing: A Journal of Practice & Theory*, **9**: 1-28.

Maignan, I. and O. C. Ferrell, 2000, “Measuring Corporate Citizenship in Two Countries: The Case of the United States and France”. *Journal of Business Ethics*, **23**: 283–297.

Marquardt, C. A. and C. I. Wiedman, 2004, “How are earnings managed? An examination of specific accruals”. *Contemporary Accounting Research*, **21:** 461-491.

McNichols, M. F., 2000, “Research design issues in earnings management studies”. *Journal of Accounting and Public Policy*, **19**: 313–345.

McNichols, M. F., 2002, “Discussion of the quality of accruals and earnings: The role of accrual estimation errors”. *The Accounting Review*, **77(SUPPL.)**: 61–69.

Perols, J. L. and B. A. Lougee, 2011, “The relation between earnings management and financial statement fraud”. *Advances in Accounting*, **27**: 39–53.

Perry, S. E. and T. H. Williams, 1994, “Earnings management preceding management buyout offers”. *Journal of Accounting and Economics*, **18:** 157–179.

Persons, O. S., 1995, “Using financial statement data to identify factors associated with fraudulent financial reporting”. *Journal of Applied Business Research*, **11**: 38-46.

Prior, D., J. Surroca and J. A. Tribó, 2008, “Are socially responsible managers really ethical? Exploring the relationship between earnings management and corporate social responsibility”. *Corporate Governance*, **16**: 160-177.

Quazi, A. M. and D. O’Brien, 2000, “An Empirical Test of a Cross-National Model of Corporate Social Responsibility”. *Journal of Business Ethics*, **25**: 33–51.

Rosner, R. L., 2003, “Earnings manipulation in failing firms”. *Contemporary Accounting Research*, **20**: 361-408.

Roychowdhury, S., 2006, “Earnings management through real activities manipulation”. *Journal of Accounting and Economics*, **42**: 335–370.

Ruf, B. M., K. Muralidhar and K. Paul, 1998, “The Development of a Systematic, Aggregate Measure of Corporate Social Performance”. *Journal of Management*, **24**: 119–133.

Shafer, W. E., 2013, “Ethical climate, social responsibility, and earnings management”. *Journal of Business Ethics*, 1-18.

Shih, K. H., C. C. Cheng and Y. H. Wang, 2011, “Financial information fraud risk warning for manufacturing industry - using logistic regression and neural network”. *Romanian Journal of Economic Forecasting*, **14**: 54-71.

Singhapakdi, A., S. J. Vitell, K. C. Rallapalli and K. L. Kraft, 1996, “The Perceived Role of Ethics and Social Responsibility: A Scale Development”. *Journal of Business Ethics*, **15**: 1131–1140.

Siregar, S. V. and S. Utama, 2008, “Type of earnings management and the effect of ownership structure, firm size, and corporate-governance practices: Evidence from Indonesia”. *The International Journal of Accounting*, **43**: 1-27.

Stubben, S. R., 2010, “Discretionary Revenues as a Measure of Earnings Management”. *The Accounting Review*, **85**: 695-717.

Summers, S. L. and J. T. Sweeney, 1998, “Fraudulent Misstated Financial Statements and Insider Trading: An Empirical Analysis”. *The Accounting Review*, **73**: 131–46.

Teoh, S. H., I. Welch and T. J. Wong, 1998, “Earnings management and the long-run underperformance of seasoned equity offerings”. *Journal of Financial Economics*, **50**: 63–99.

Turker, D., 2009, “Measuring corporate social responsibility: A scale development study”. *Journal of Business Ethics*, **85**: 411-427.

Waddock, S. A., 2003, “Myths and realities of social investing”. *Organization & Environment*, **16**: 369–380.

Williams, C. C., 2009, "Illegitimate wage practices in central and eastern Europe: A study of the prevalence and impacts of "envelope wages"". *Debatte*, **17**: 65-83.

Wongsunwai, W., 2013, "The effect of external monitoring on accrual-based and real earnings management: Evidence from venture-backed initial public offerings". *Contemporary Accounting Research*, **30**: 296-324.

Zang, A., 2012, "Evidence on the trade-off between real activities manipulation and accrual-based earnings management". *The Accounting Review*, **87**: 675–703.

Table 1: Variable definitions

Variable	Description	Pred. Sign	Calculation
EARNINGS MANAGEMENT			
Aggregate accrual-based:			
DAC	Discretionary accruals	?	Residuals of modified Jones model (Dechow *et al*., 1995) with additional control for firm performance (Kothari *et al*., 2005) (Eq. 1)
ABSDAC	Absolute value of discretionary accruals	+	Absolute value of *DAC*
Revenue accrual-based:			
DREV	Discretionary revenue accruals	+	Residuals from Caylor's (2010) model (Eq. 5)
ABSDREV	Absolute value of discretionary revenue accruals	+	Absolute value of *DREV*
Expense accrual-based:			
DEXP	Discretionary expense accruals	+	Residuals from Eq. (6)
ABSDEXP	Absolute value of discretionary expense accruals	+	Absolute value of *DEXP*
RM:			
ABMAT	Abnornal material expenses	+	Residuals from Eq. (7)

Table 1: Variable definitions

Variable	Description	Pred. Sign	Calculation
ABPER	Abnormal personnel expenses	–	Residuals from Eq. (7)
ABSERV	Abnormal service expenses	–	Residuals from Eq. (8)
ABCFO	Abnormal CFO	?	Residuals from Eq. (9) (Roychowdhury, 2006)
Unadjusted EM proxies:			
ACCR	Total accruals deflated by lagged total assets	?	Total accruals Eq. (3)/total assets$_{t-1}$
CH_REC	Change in receivables deflated by lagged total assets	+	(Receivables$_t$ - receivables$_{t-1}$)/total assets$_{t-1}$
CH_INV	Change in inventory deflated by lagged total assets	?	(Inventory$_t$ - inventory$_{t-1}$)/total assets$_{t-1}$
CH_PAY	Change in payables deflated by lagged total assets	+	(Payables$_t$ - payables$_{t-1}$)/total assets$_{t-1}$
PERTA	Personnel expenses to lagged total assets	–	Personnel expenses/total assets$_{t-1}$
MATTA	Material expenses to lagged total assets	+	Material expenses/total assets$_{t-1}$
SERVTA	Service expenses to lagged total assets	–	Service expenses/total assets$_{t-1}$
ASSET COMPOSITION:			
INTA	Intangible assets to total assets	?	Intangible assets/total assets
CATA	Current assets to total assets	?	Current assets/total assets
RECTA	Receivables to total assets	+	Receivables/total assets
INVTA	Inventory to total assets	–	Inventory/total assets
PERFORMANCE:			
ROA	Return on assets	–	Earnings before interests and extraordinary items/total assets
ABS_CH_ROA	Absolute value of change in ROA	+	Absolute value of: ROA$_t$-ROA$_{t-1}$
REVTA	Revenue to assets	–	Revenue$_t$/total assets$_{t-1}$
SERVREV	Service expenses to sales	–	Service expenses/sales
MATREV	Material expenses to sales	+	Material expenses/sales
PERSREV	Personnel expenses to sales	–	Personnel expenses/sales
ABS_CH_PERSREV	Absolute value of change in personnel expenses over sales	+	Absolute value of: (Personnel expenses/sales)$_t$ - (Personnel expenses/sales)$_{t-1}$
ABS_CH_MATREV	Absolute value of change in material expenses over sales	+	Absolute value of: (material expenses/sales)$_t$ - (material expenses/sales)$_{t-1}$
ABS_CH_SERVREV	Absolute value of change in service expenses over sales	+	Absolute value of: (service expenses/sales)$_t$ - (service expenses/sales)$_{t-1}$

Table 1: Variable definitions

Variable	Description	Pred. Sign	Calculation
ABS_CH_NI	Absolute value of change in net income	+	Absolute value of: (net income$_t$-net income$_{t-1}$)/total assets$_{t-1}$
ABS_CH_CFO	Absolute value of change in CFO	+	Absolute value of: (CFO$_t$-CFO$_{t-1}$)/total assets$_{t-1}$
ABS_CH_RECREV	Absolute value of change in receivables to sales	+	Absolute value of: (receivables/sales)$_t$-(receivables/sales)$_{t-1}$
ABS_CH_PAYEXP	Absolute value of change in payables to purchases	+	Absolute value of: (payables/expenses)$_t$-(payables/expenses)$_{t-1}$
DEBT:			
LEV	Leverage	+	Total liabilities/total assets
LEVBANK	Bank indebtedness	–	Bank debts/total assets
LIQUIDITY:			
CRATIO	Current ratio	–	Current assets/current liabilities
ABS_CH_CRATIO	Absolute value of change in current ratio	+	Absolute value of: $CRATIO_t$ - $CRATIO_{t-1}$
GROWTH:			
GROWTH	Percentage change in total assets	+	(Total assets$_t$-total assets$_{t-1}$)/total assets$_{t-1}$
NON-FINANCIAL:			
PERSEMPL	Personnel expenses to employees	–	Personnel expenses/number of employees
REVEMPL	Revenue to employee	+	Revenue$_t$/employees$_{t-1}$
DIF_GROWTH_EMPL	Percentage change in total assets less percentage change in number of employees	–	GROWTH-(employees$_t$-employees$_{t-1}$)/employees$_{t-1}$
YEAR	Fiscal year	?	Dummy variables representing the fiscal year
IND	Industry	?	Dummy variables representing industry defined by the two-digit SIC code

Table 2: Industry distribution of LMFs and AIDA population of active unlisted firms with available financial data from 2003 to 2012 restricted to LMFs industries (LWFs)

Sic code	Industry description	AIDA population		LMFs	
		Freq.	%	Freq.	%
01	Agricultural production-crops	644	0.82%	4	2.02%
14	Mining and quarrying of nonmetallic minerals, except fuels	463	0.59%	8	4.04%
15	Building construction-general contractors and operative builders	5,486	7.00%	33	16.67%
16	Heavy construction other than building construction-contractors	524	0.67%	3	1.52%
17	Construction-special trade contractors	4,032	5.15%	8	4.04%
20	Food and kindred products	3,224	4.12%	6	3.03%
25	Furniture and fixtures manufacturing	829	1.06%	3	1.52%
28	Chemicals and allied products manufacturing	1,598	2.04%	1	0.51%
29	Petroleum refining and related industries	158	0.20%	2	1.01%
32	Stone, clay, glass and concrete products manufacturing	1,960	2.50%	11	5.56%
34	Fabricated metal products, except machinery and transportation equipment	7,038	8.98%	1	0.51%
42	Motor freight transportation and warehousing	2,894	3.69%	18	9.09%
44	Water transportation	586	0.75%	1	0.51%
45	Transportation by air	95	0.12%	1	0.51%

47	Transportation services	1,884	2.40%	3	1.52%
49	Electric, gas and sanitary services	1,419	1.81%	6	3.03%
50	Wholesale trade, durable goods	14,064	17.95%	22	11.11%
51	Wholesale trade, nondurable goods wholesale dealing in	7,821	9.98%	17	8.59%
52	Building materials, hardware, garden supply, and mobile home dealers wholesale dealing in	1,018	1.30%	1	0.51%
53	General merchandise stores	324	0.41%	1	0.51%
54	Food stores	1,737	2.22%	15	7.58%
55	Automotive dealers and gasoline service stations	536	0.68%	2	1.01%
56	Apparel and accessory stores	1,920	2.45%	2	1.01%
57	Home furniture, furnishings, and equipment stores	872	1.11%	1	0.51%
58	Eating and drinking places	1,007	1.29%	2	1.01%
59	Miscellaneous retail	1,475	1.88%	1	0.51%
65	Real estate	2,239	2.86%	6	3.03%
70	Hotels, rooming houses, camps, and other lodging places	1,600	2.04%	3	1.52%
72	Personal services	327	0.42%	1	0.51%
73	Business services	5,001	6.38%	2	1.01%
75	Automotive repair, services, and parking	882	1.13%	1	0.51%
79	Amusement and recreation services	744	0.95%	4	2.02%
80	Health services	1,165	1.49%	5	2.53%
81	Legal services	19	0.02%	1	0.51%
87	Engineering, accounting, research, management, and related services	2,755	3.52%	2	1.01%
Total		**78,340**	**100%**	**198**	**100%**

Source: AIDA database, 2013.

Table 3: Sample selection

	Number of firms
LMFs sample	
LMFs definitively confiscated at November 5th 2012 provided by ANBSC	1,663
Less: LMFs provided by ANBSC with data unavailable on AIDA database	-1,609
Add: LMFs found in AIDA database with status confiscated	118
Add: confiscated LMFs found in online newspapers with data available on AIDA	52
Less: LMFs confiscated before 2005 with pre-confiscation data unavailable on AIDA	-26
Final LMFs sample	**198**
LMF-years in detection model	**426**
LWFs control sample	
Aida population of active and unlisted firms with available financial data from 2003 to 2012 in the same two-digit SIC industries as LMFs	78,340
Less: LWFs not matched to LMFs by year, sector and size	-77,922
Final LWFs in detection model	**418**
LWF-years in detection model	**426**

Source: ANBSC and AIDA database, 2013.

Table 4: LMFs by confiscation year

Confiscation year	Number of confiscated LMFs	Percentage
2005	1	0.51%
2006	9	4.55%
2007	18	9.09%
2008	24	12.12%
2009	19	9.60%
2010	24	12.12%
2011	35	17.68%
2012	37	18.69%
2013	31	15.66%
Total	**198**	**100.00%**

Source: ANBSC and AIDA database, 2013.

Table 5: Descriptive statistics and pairwise variable comparison between LMFs and LWFs									
		LMFs		LWFs			Difference (LMFs - LWFs)		
Variable	**N**	**Mean**	**Median**	**Mean**	**Median**	**Pred. Sign**	**Mean**	**Median**	**Test**
EARNINGS MANAGEMENT									
Aggregate accrual-based:									
DAC	516	0.012	-0.004	-0.006	0.003	?	0.018	-0.007	
ABSDAC	516	0.188	0.115	0.145	0.090	+	0.043	0.026	***
Revenue accrual-based:									
DREV	460	0.031	0.016	0.024	-0.005	+	0.007	0.020	
ABSDREV	460	0.146	0.089	0.115	0.064	+	0.032	0.026	***
Expense accrual-based:									
DEXP	478	0.028	0.009	0.001	-0.007	+	0.026	0.017	**
ABSDEXP	478	0.146	0.091	0.108	0.062	+	0.038	0.028	***
RM:									
ABMAT	601	0.107	0.061	-0.018	-0.017	+	0.125	0.078	***
ABPER	601	-0.024	-0.047	-0.012	-0.024	–	-0.012	-0.023	**
ABSERV	741	-0.012	-0.072	0.005	-0.035	–	-0.017	-0.037	***
ABCFO	543	-0.026	-0.005	0.013	-0.002	?	-0.039	-0.002	**
Unadjusted EM proxies:									
ACCR	543	0.012	-0.020	-0.015	-0.017	?	0.028	-0.003	

Table 5: Descriptive statistics and pairwise variable comparison between LMFs and LWFs

Variable	N	LMFs Mean	LMFs Median	LWFs Mean	LWFs Median	Pred. Sign	Difference (LMFs - LWFs) Mean	Difference (LMFs - LWFs) Median	Test
CH_REC	625	0.097	0.025	0.053	0.013	+	0.044	0.012	***
CH_INV	741	0.039	0.000	0.015	0.000	?	0.024	0.000	***
CH_PAY	552	0.075	0.014	0.040	0.013	+	0.034	0.001	**
PERTA	741	0.202	0.108	0.214	0.144	–	-0.012	-0.036	***
MATTA	741	0.945	0.442	0.724	0.380	+	0.221	0.063	***
SERVTA	741	0.391	0.182	0.437	0.254	–	-0.046	-0.073	***
ASSET COMPOSITION									
INTA	967	0.035	0.004	0.025	0.004	?	0.010	0.000	*
CATA	966	0.743	0.819	0.734	0.807	?	0.010	0.012	*
RECTA	875	0.389	0.380	0.374	0.356	+	0.015	0.024	**
INVTA	967	0.184	0.054	0.185	0.097	–	-0.001	-0.044	**
PERFORMANCE									
ROA	967	0.040	0.035	0.059	0.041	–	-0.018	-0.007	***
ABS_CH_ROA	741	0.051	0.024	0.041	0.021	+	0.010	0.003	**
REVTA	741	1.585	1.041	1.503	1.165	–	0.082	-0.124	
SERVREV	908	0.292	0.176	0.340	0.266	–	-0.048	-0.090	***
MATREV	908	0.550	0.564	0.440	0.434	+	0.110	0.130	***
PERSREV	908	0.171	0.108	0.168	0.124	–	0.003	-0.016	**
ABS_CH_PERSREV	684	0.057	0.024	0.033	0.013	+	0.023	0.011	***
ABS_CH_MATREV	684	0.153	0.050	0.087	0.028	+	0.066	0.022	***
ABS_CH_SERVREV	684	0.121	0.033	0.083	0.026	+	0.038	0.007	***
ABS_CH_NI	741	0.046	0.013	0.032	0.013	+	0.014	0.000	***
ABS_CH_CFO	363	0.265	0.137	0.221	0.129	+	0.043	0.008	*
ABS_CH_RECREV	571	0.299	0.111	0.169	0.056	+	0.131	0.054	***
ABS_CH_PAYEXP	547	0.357	0.137	0.204	0.065	+	0.153	0.072	***
DEBT									
LEV	967	0.774	0.840	0.684	0.736	+	0.090	0.103	***
LEVBANK	807	0.134	0.046	0.164	0.100	–	-0.030	-0.054	***
LIQUIDITY									
CRATIO	962	1.365	1.054	1.457	1.175	–	-0.092	-0.122	***
ABS_CH_CRATIO	734	0.401	0.118	0.317	0.102	+	0.084	0.015	
GROWTH									
GROWTH	741	0.242	0.110	0.102	0.036	+	0.140	0.074	***
NON-FINANCIAL									
PERSEMPL	908	27.251	26.373	34.192	32.173	–	-6.941	-5.800	***
REVEMPL	703	781.379	280.656	533.141	274.466	+	248.238	6.190	
DIF_GROWTH_EMPL	697	-0.114	-0.014	0.037	0.033	–	-0.151	-0.048	***

Notes: The sample full period spans 2003–2012. $*$, $**$ and $***$ denote significance levels at 10%, 5% and 1%, respectively, based on a two-tailed Wilcoxon signed-rank test for the differences in medians between paired samples. See Table 1 for variable definitions. We apply non parametric Wilcoxon signed-rank test rather than Student's t-test for differences in means given that untabulated Skewness/Kurtosis tests for Normality show non-normality of most of the variables. However, both tests mostly perform the same.

.

Table 6: Pearson correlations between EM related variables

	DAC	*ABSDAC*	*DREV*	*ABSDREV*	*DEXP*	*ABSDEXP*	*ABMAT*	*ABPER*	*ABSERV*	*ABCFO*	*ACCR*	*CH_REC*	*CH_PAY*	*CH_INV*	*PERTA*	*MATTA*	*SERVTA*
DAC	1																
ABSDAC	0.1***	1															
DREV	0.0	0.1***	1														
ABSDREV	0.0	0.4***	0.4***	1													
DEXP	-0.1***	0.1***	0.6***	0.3***	1												
ABSDEXP	0.0	0.4***	0.2***	0.5***	0.3***	1											
ABMAT	0.1***	0.1**	0.0	0.0	0.1***	0.0	1										
ABPER	0.0	0.0	0.0	0.0	0.0	-0.1***	-0.2***	1									
ABSERV	0.1*	0.2***	0.1***	0.2***	0.2***	0.2***	-0.5***	-0.2***	1								
ABCFO	-0.9***	-0.1***	0.0	0.0	0.1***	-0.1**	-0.2***	0.0	0.0	1							
ACCR	0.9***	0.2***	0.0	0.0	-0.1***	0.0	0.1***	0.0	0.0	-0.9***	1						
CH_REC	0.0	0.2***	0.9***	0.5***	0.5***	0.3***	0.1*	0.0	0.3***	0.0	0.0	1					
CH_PAY	-0.3***	0.2***	0.5***	0.4***	0.8***	0.3***	0.1**	0.0	0.3***	0.3***	-0.4***	0.7***	1				
CH_INV	0.4***	0.3***	0.0	0.1**	0.3***	0.1***	0.2***	0.1***	0.1***	-0.4***	0.5***	0.0	0.2***	1			
PERTA	0.0	0.1***	0.1***	0.2***	0.0	0.1***	-0.2***	0.7***	0.0	0.0	-0.1**	0.2***	0.1***	0.0	1		
MATTA	0.0	0.1***	0.1***	0.2***	0.1***	0.2***	0.3***	-0.1**	0.0	0.0	0.0	0.2***	0.2***	0.2***	0.1***	1	
SERVTA	0.0	0.3***	0.1***	0.4***	0.1***	0.3***	-0.5***	-0.1***	0.7***	0.0	0.0	0.2***	0.2***	0.1***	0.3***	0.0	1

Notes: *, ** and *** denote significance levels at 10%, 5% and 1%, respectively, based on a two-tailed test. See Table 1 for variable definitions.

Table 7: Logistic regression comparing LMFs with LWFs			
Variable	**Pred. Sign**	**Estimate**	**p-value**
CH_REC	+	-1.249	0.024
CH_INV	?	-2.618	0.015
ABMAT	+	0.991	0.009
PERTA	–	-0.610	0.249
INTA	?	4.640	0.002
CATA	?	-0.630	0.382
RECTA	+	0.838	0.199
INVTA	–	-2.101	0.015
REVTA	–	-0.335	0.003
SERVREV	–	-1.383	0.008
ABS_CH_PERSREV	+	4.931	0.005
ABS_CH_MATREV	+	0.849	0.183
ABS_CH_NI	+	1.361	0.400
LEV	+	4.383	0.000
LEVBANK	–	-3.554	0.000
GROWTH	+	1.887	0.000
PERSEMPL	–	-0.023	0.000
REVEMPL	+	0.000	0.030
DIF_GROWTH_EMPL	–	-0.570	0.003
IND dummies		Yes	
YEAR dummies		Yes	
Intercept		-0.604	0.493
Number of observations		852	
LR chi2(57)		271.89	0.000
Pseudo R2		0.230	
Area under ROC Curve		0.816	
Correctly classified (cut-off = 0.50)			
LMFs		76.29%	
LWFs		76.53%	
Overall		76.41%	

Notes: The p-values are two-tailed. See Table 1 for variable definitions.

Table 8: Illicit activities and related reflecting variables of the detection model	
Illicit activity	**Reflecting variables**
Fraudulent accounting manipulations	*CH_REC; CH_INV; ABMAT; INTA; CATA; RECTA; INVTA; REVTA; SERVREV; ABS_CH_MATREV; ABS_CH_NI; GROWTH*
Money laundering through fictitious transactions	*CH_REC; ABMAT; CATA; RECTA; REVTA; SERVREV; ABS_CH_MATREV; ABS_CH_NI; LEV; LEVBANK; GROWTH; DIF_GROWTH_EMPL*
Income tax/ value added tax evasion	*CH_REC; CH_INV; ABMAT; INVTA; REVTA; ABS_CH_MATREV; ABS_CH_NI*
Wage compression including evasion of social security contributions	*PERTA; ABS_CH_PERSREV; PERSEMPL; REVEMPL; DIF_GROWTH_EMPL*
Supplier intimidation	*LEV*

Notes: See Table 1 for variable definitions.

Table 9: Logistic regressions excluding hold-out samples						
	2013 excluded		2012 excluded		2011 excluded	
Variable	**Estimate**	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**
CH_REC	-1.532	0.022	-1.460	0.017	-1.579	0.012
CH_INV	-1.855	0.167	-1.836	0.137	-2.691	0.024
ABMAT	0.887	0.060	1.173	0.007	0.770	0.074
PERTA	-2.542	0.000	-0.261	0.675	-0.186	0.749
INTA	2.210	0.187	5.350	0.002	5.762	0.001
CATA	-0.575	0.527	-0.775	0.349	0.182	0.817
RECTA	0.347	0.683	0.996	0.164	0.811	0.255
INVTA	-3.430	0.002	-2.063	0.030	-1.995	0.036
REVTA	-0.164	0.257	-0.236	0.054	-0.300	0.009
SERVREV	-2.028	0.001	-0.830	0.178	-1.326	0.023
ABS_CH_PERSREV	3.079	0.154	6.520	0.001	3.518	0.079
ABS_CH_MATREV	1.661	0.052	-0.028	0.968	1.729	0.016
ABS_CH_NI	3.378	0.097	0.196	0.917	1.417	0.422
LEV	4.589	0.000	4.409	0.000	3.602	0.000
LEVBANK	-4.182	0.000	-3.091	0.000	-3.530	0.000
GROWTH	2.112	0.000	1.779	0.001	1.836	0.000
PERSEMPL	-0.025	0.000	-0.025	0.000	-0.022	0.000
REVEMPL	0.000	0.105	0.000	0.165	0.000	0.016
DIF_GROWTH_EMPL	-0.842	0.001	-0.555	0.009	-0.540	0.010
IND dummies	Yes		Yes		Yes	
YEAR dummies	Yes		Yes		Yes	
Intercept	0.098	0.927	-0.799	0.393	-1.173	0.267
Number of observations	632		648		680	
LR chi2	221.8	0.000	207.9	0.000	196.51	0.000
Pseudo R2	0.253		0.231		0.2085	
Area under ROC Curve	0.827		0.816		0.806	
Correctly classified hold-out samples (cut-off = 0.50):						
Confiscation year	**2013**		**2012**		**2011**	
Number of observations	220		204		172	
LMFs	66.36%		75.49%		81.40%	
LWFs	72.73%		67.65%		74.42%	
Overall	69.55%		71.57%		77.91%	

Notes: The p-values are two-tailed. See Table 1 for variable definitions.

Table 10: Industry distribution of undetected and detected LMFs					
Sic code	**Industry description**	**Undetected LMFs**		**Detected LMFs**	
		Freq.	**%**	**Freq.**	**%**
01	Agricultural production-crops	3	2.97%	4	1.23%
14	Mining and quarrying of nonmetallic minerals, except fuels	5	4.95%	13	4.00%
15	Building construction-general contractors and operative builders	21	20.79%	50	15.38%
16	Heavy construction other than building construction-contractors	2	1.98%	4	1.23%
17	Construction-special trade contractors	3	2.97%	22	6.77%
20	Food and kindred products	2	1.98%	5	1.54%
25	Furniture and fixtures manufacturing	2	1.98%	11	3.38%
28	Chemicals and allied products manufacturing	0	0.00%	1	0.31%
29	Petroleum refining and related industries	1	0.99%	6	1.85%
32	Stone, clay, glass and concrete products manufacturing	7	6.93%	21	6.46%
34	Fabricated metal products, except machinery and transportation equipment	0	0.00%	5	1.54%
42	Motor freight transportation and warehousing	11	10.89%	32	9.85%
44	Water transportation	2	1.98%	4	1.23%
47	Transportation services	1	0.99%	2	0.62%
49	Electric, gas and sanitary services	5	4.95%	9	2.77%
50	Wholesale trade, durable goods	16	15.84%	47	14.46%
51	Wholesale trade, nondurable goods wholesale dealing in	9	8.91%	27	8.31%
53	General merchandise stores	0	0.00%	1	0.31%
54	Food stores	4	3.96%	10	3.08%
55	Automotive dealers and gasoline service stations	0	0.00%	6	1.85%
56	Apparel and accessory stores	0	0.00%	4	1.23%
59	Miscellaneous retail	0	0.00%	2	0.62%
65	Real estate	0	0.00%	5	1.54%
70	Hotels, rooming houses, camps, and other lodging places	0	0.00%	5	1.54%
72	Personal services	1	0.99%	2	0.62%
73	Business services	2	1.98%	3	0.92%
75	Automotive repair, services, and parking	0	0.00%	6	1.85%
79	Amusement and recreation services	0	0.00%	1	0.31%
80	Health services	2	1.98%	8	2.46%
81	Legal services	1	0.99%	6	1.85%
87	Engineering, accounting, research, management, and related services	1	0.99%	3	0.92%
Total		**101**	**100.00%**	**325**	**100.00%**

Table 11: Comparison of variables between undetected and detected LMFs									
	Undetected LMFs			Detected LMFs			Difference (Undetected - Detected)		
Variable	**N**	**Mean**	**Median**	**N**	**Mean**	**Median**	**Mean**	**Median**	**Test**
Total assets (logarithm)	101	8.508	8.493	325	8.216	8.222	0.291	0.271	**
Number employees	101	30.943	13.000	325	23.455	11.000	7.488	2.000	**
CH_REC	101	0.049	0.024	325	0.084	0.034	-0.035	-0.010	
CH_INV	101	0.025	0.000	325	0.006	0.000	0.019	0.000	**
ABMAT	101	-0.023	-0.011	325	0.157	0.111	-0.180	-0.122	***
PERTA	101	0.212	0.101	325	0.190	0.106	0.023	-0.005	
INTA	101	0.015	0.001	325	0.048	0.004	-0.032	-0.003	***
CATA	101	0.759	0.867	325	0.725	0.766	0.034	0.100	**
RECTA	101	0.464	0.469	325	0.452	0.476	0.012	-0.006	
INVTA	101	0.170	0.050	325	0.140	0.042	0.031	0.008	
REVTA	101	1.338	1.042	325	1.546	1.043	-0.209	-0.001	
SERVREV	101	0.362	0.226	325	0.243	0.169	0.119	0.057	***
ABS_CH_PERSREV	101	0.039	0.018	325	0.054	0.024	-0.015	-0.005	
ABS_CH_MATREV	101	0.081	0.022	325	0.149	0.056	-0.068	-0.034	***
ABS_CH_NI	101	0.030	0.013	325	0.040	0.012	-0.010	0.001	
LEV	101	0.679	0.709	325	0.794	0.841	-0.115	-0.132	***
LEVBANK	101	0.192	0.166	325	0.147	0.075	0.045	0.091	***
GROWTH	101	0.084	0.065	325	0.189	0.087	-0.104	-0.022	
PERSEMPL	101	38.732	30.529	325	27.641	26.343	11.092	4.187	***
REVEMPL	101	522.278	253.141	325	848.794	293.083	-326.516	-39.942	
DIF_GROWTH_EMPL	101	0.080	0.029	325	-0.164	-0.063	0.244	0.092	***

Notes: $*$, $**$ and $***$ denote significance levels at 10%, 5% and 1%, respectively, based on a two-tailed Mann–Whitney–Wilcoxon test for the differences in medians. See Table 1 for variable definitions.

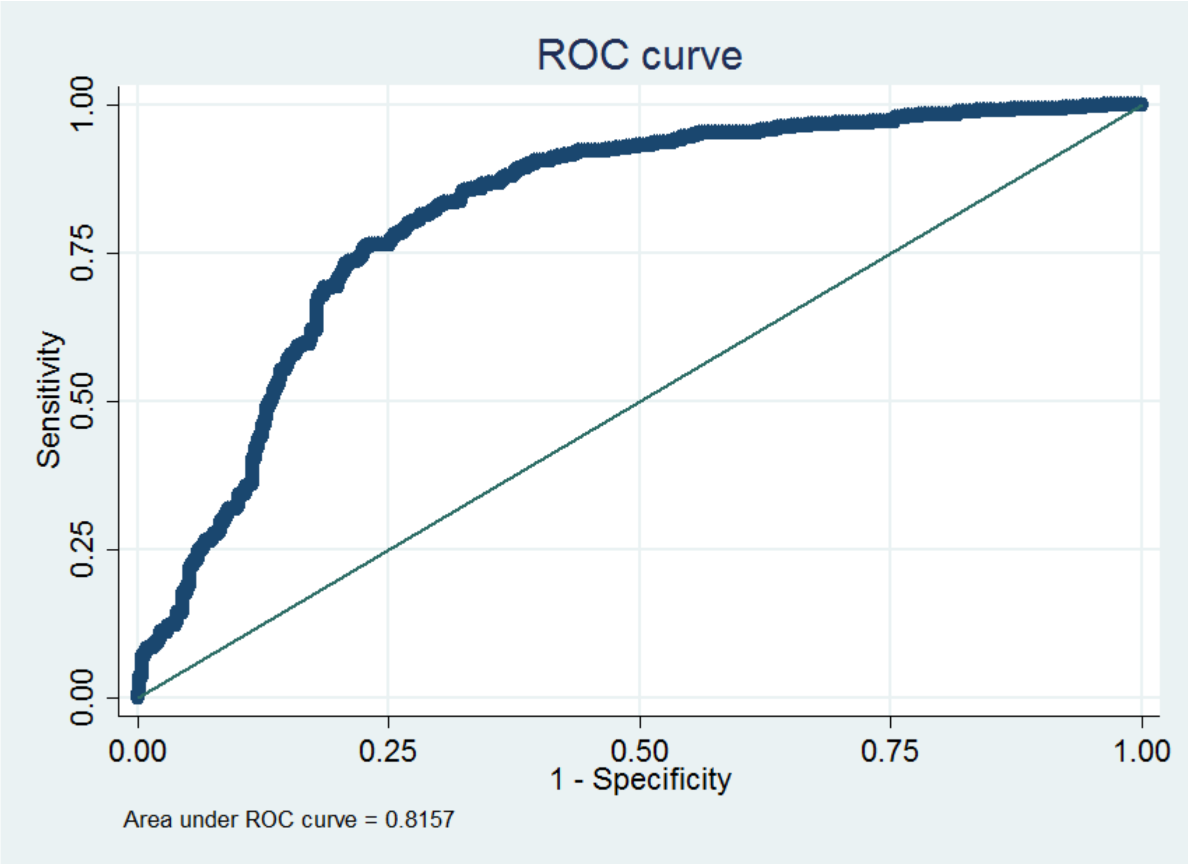

Fig.1.This figure shows the receiver operating characteristic (ROC) curve for the logistic regression results of Table 7. The ROC curve shows the relation between the proportion of LMFs detected and the proportion of false positives for all possible classification probability cut-off points.

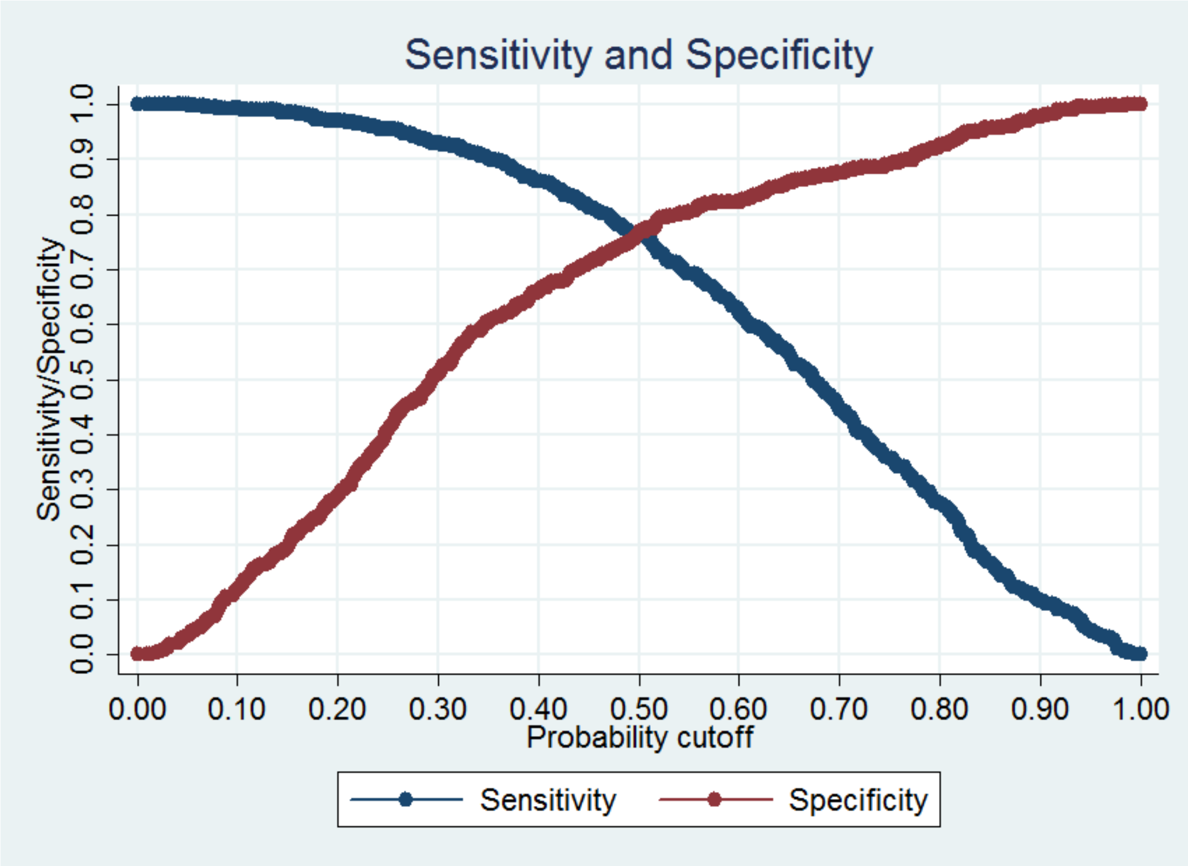

Fig.2.This figure shows the graph of sensitivity and specificity versus probability cutoff-points.