



Are Bankers “Crying Wolf”? Type I, Type II Errors and Deterrence in Anti-Money Laundering: The Italian Case

Lucia dalla Pellegrina¹ · Giorgio Di Maio² · Donato Masciandaro³ · Margherita Saraceno⁴

Received: 21 March 2021 / Accepted: 11 April 2022
© The Author(s) 2022

Abstract

Excessive and useless reporting, called the “crying wolf effect,” is a crucial shortcoming that any anti-money laundering (AML) design aims to address. For this reason, in recent years, AML policies in both the US and Europe have switched from a rule-based to a risk-based approach. This study theoretically and empirically investigates whether the risk-based approach delivers the expected results. The theoretical model shows that a trade-off can emerge between accuracy (fewer type-I and type-II errors) and deterrence. The empirical analysis, conducted after the risk-based approach was introduced in Italy, confirms this trade-off. More specifically, deterrence seems a priority, whereas accuracy is sacrificed. In this respect, the data suggest that Italian bankers are likely to “cry wolf.”

Keywords Anti-money laundering · Suspicious transaction reporting · Reporting test · Type-I error · Type-II error · Deterrence · Italy

JEL Classification G2 · K2 · K4

✉ Donato Masciandaro
donato.masciandaro@unibocconi.it

Lucia dalla Pellegrina
lucia.dallapellegrina@unimib.it

Giorgio Di Maio
giorgiodimaio@gmail.com

Margherita Saraceno
margherita.saraceno@unipv.it

¹ Department of Economics, Management and Statistics and CISEPS, University of Milano-Bicocca, Milan, Italy

² Department of Economics, University of Insubria, Varese, Italy

³ Department of Economics, Bocconi University and SUERF, Milan, Italy

⁴ Department of Law, University of Pavia, Pavia, Italy

1 Introduction

Excessive and useless reporting, known as the “crying wolf effect” (Takats 2011), is a crucial shortcoming that any anti-money laundering (AML) design aims to address and fix. The “crying wolf effect” harms the informational value of reports that banks and other professionals are obliged to file to comply with AML regulations.

The AML system in Europe and in the United States consists of a three-layer hierarchy of enforcers: financial intermediaries and other professionals; a Financial Intelligence Unit (FIU), normally established at the central bank; and the judiciary system. At the first level, financial intermediaries and other professionals are required to monitor all financial transactions and report suspected acts of money laundering to the FIU by filing a suspicious transaction report (STR).

Initially, the AML system followed a rule-based approach. Financial intermediaries and other professionals used a set of standardized criteria (determined by the law and the FIU) to identify suspicious transactions and report them to the FIU. In that system, the role of financial intermediaries and other professionals was relatively passive. A chief problem of the rule-based approach was the high number of STRs erroneously issued by financial intermediaries and professionals. The high incidence of type-I errors (false positives) in the rule-based AML system was considered inefficient because it wasted the FIU’s resources. Moreover, it was ineffective in deterring money laundering and detrimental for intermediaries and professionals (especially from a reputational perspective). However, simply raising the bar by imposing stricter rules and criteria for reporting a transaction as suspicious to the FIU was not a solution. In this regard, false negatives represent an additional problem for AML systems, as decisions to not report potential money-laundering transactions (type-II errors) both dilute deterrence and make the financial system less reliable (Demetis 2010).

Between 2007 and 2010, AML policies in both the US and Europe switched from a rule-based reporting system to a risk-based system in which all layers of the system need to respond to money-laundering threats in ways that are proportionate to the risks involved.¹ In particular, financial intermediaries and other professionals are required to play an active role in identifying suspicious transactions (Black and Baldwin 2010; Dalla Pellegrina and Masciandaro 2009). They must exploit their knowledge and other information regarding the financial habits of their customers [the know-your-customer (KYC) approach] to better determine which transactions should be reported as suspicious to the FIU. They must also apply their subjective judgment to assess the actual risk of a transaction being money laundering. In fact, intermediaries and professionals are required to adjust their reporting criteria and, therefore, move up or down their decisional bars (reporting test) when deciding whether to report a transaction to the FIU based on the actual risk of money laundering (Axelrod 2017; Lowe 2017).

The risk-based approach was introduced starting from 2007 mainly to avoid over-reporting to the FIU without allowing type-II errors to explode. In general, the risk-based AML system aims to increase the reliability and accuracy of the STRs that financial intermediaries and other professionals send to the FIU. In this vein, the KYC approach (see Jeans 2016) should allow the reduction of the number of both type-I

¹ https://www.anti-moneylaundering.org/EU_Chart.aspx

and type-II errors at the first level of the AML system. At the second level of the AML system, the FIU analyses the collected STRs and reports the transactions that it deems to be money-laundering acts to the judicial authority. Type-I errors in STRs submitted by financial intermediaries and professionals are typically dismissed by the FIU because they are not considered true money-laundering transactions.² The judicial authority, the third level of the AML system, collects reports from the FIU and decides whether to issue a referral to trial.

Type-II errors committed on both the first and second levels of the AML system can be detected by the judicial authority. The latter also collects reports on money laundering that come from institutions other than the FIU and from actors other than financial intermediaries and professionals. For instance, money laundering can be detected by the police while investigating other crimes. Sometimes, criminal organizations' confessions describe how illegal funds are laundered and how those activities avoid AML measures (Arnone and Borlini 2010; Barone and Masciandaro 2019). Although they are possible, errors at the level of the law enforcement system are not included in this analysis, which focuses on the first level of the AML system.

As Unger and van Waarden (2009) discussed, despite the aim of making the (first level) of the AML system more reliable, the impact of the risk-based approach differs across countries. In some countries, over-reporting decreased, and the overall quality of the reported information improved. However, this is not the case in other countries.³

This study aims to investigate whether the risk-based approach introduced in Italy in 2009 had the expected results in terms of increased reporting accuracy and, in particular, a lower rate of type-I errors at the first level of the AML system. The analysis is based on a theoretical model that describes the relations between the reporting test, type-I errors, type-II errors, and their sum (a measure of accuracy), and the deterrence of money laundering activities. In general, the empirical aim is to test the most important implications of the theoretical model using data from the Italian FIU in the aftermath of the risk-based approach introduction (2009–2012).⁴ We perform regression analysis supported by an approach based on the concept of sufficient statistics (Chetty 2009). Our results show that deterrence is prioritized, although an increase in type-I errors likely occurred following the introduction of the risk-based system. We thereby make inferences regarding the trend in type-II errors using the predictions of the theoretical model. We conclude that type-II errors decreased during the period of interest.

² The FIU can commit both type-I and type-II errors as well (i.e., filing a report when the transaction is not a money-laundering act and not filing a law report when the transaction is a money-laundering act). However, we focus on the efficiency of the first level of the AML system. As such, we assume that the second-level authorities have perfect foresight.

³ In particular, Unger and van Waarden (2009) show that over-reporting decreased and quality increased in the Netherlands. In other countries, including the US, over-reporting increased with detrimental effects on quality (the “crying wolf effect”). However, according to Gara et al. (2019), reporting activity in Italy increased without reduction in the quality of information. In fact, the quality of information improved.

⁴ We do not go further with the observation period because important innovations in the processing system have been introduced at the FIU. These made it possible to considerably reduce the stock of reports created immediately after the introduction of the risk-based approach (FIU, Annual Reports, 2011–2012). Moreover, subsequent amendments to the regulations, made in subsequent years, would have altered the verification of the hypotheses of the model proposed here.

The remainder of this paper proceeds as follows. In the next section, we present our theoretical framework. In Sect. 3, we empirically assess the model and discuss the policy implications. Section 4 concludes this paper. The three appendices provide all details of the analysis.

2 The Model

All three levels of the AML system make their decisions based on pieces of evidence that support or contradict the hypothesis that a certain transaction involves money laundering. In particular, financial intermediaries and professionals decide whether to issue an STR for a given transaction mainly by considering the transaction attributes and customer characteristics (Gara and Pauselli 2015).

The professional assessment of intermediaries is relevant and must be considered together with all the other elements when assessing the risk that a transaction is money laundering. Financial intermediaries and professionals issue an STR when the evidence corroborating the idea that a transaction involves money laundering is greater than a certain threshold they set to conduct the anti-money laundering activity (*reporting test*).

2.1 Evidence of Money Laundering

We assume that each attribute of a transaction, including judgements on the parties involved in the transaction, is a piece of evidence that takes either a positive sign when consistent with the suspicion of a money laundering or a negative sign when against that suspicion. We do not consider how the net evidence is produced; however, all pieces of evidence combine in the *net evidence*, which is the continuous random variable E (see Rizzolli and Saraceno (2013)). In particular, we assume the following:

- $(E|Guilty) \sim N(\mu_G, \sigma_G^2)$; we define $g_E(e)$ as the probability density function of E conditional on the client being guilty and $G_E(e)$ as its cumulative distribution function.
- $(E|Innocent) \sim N(\mu_I, \sigma_I^2)$; we define $i_E(e)$ as the probability density function of E conditional on the client being innocent and $I_E(e)$ as its cumulative distribution function.
- $\mu_I < 0 < \mu_G$, and $\sigma_I^2 = \sigma_G^2$ (first-order stochastic dominance).
- $i_E(-e) = g_E(e)$ (symmetry).

The assumption of normality is for the sake of simplicity and is consistent with the graphic illustration of the model (Figs. 1, 2, 3, 4, 5, 6, 7); it might be relaxed. Regarding the assumption of first-order stochastic dominance of G_E over I_E , it is consistent with the idea that the net evidence is informational to some extent: on average, the net evidence is positive for a money-laundering transaction and negative for a transaction that does not involve money laundering. As the observed net evidence e increases, it becomes less likely that it is about an innocent client. Conversely, as the observed net evidence decreases, it is less likely to concern a guilty client.

Fig. 1 Reporting test and probabilities of a correct or incorrect reporting decision

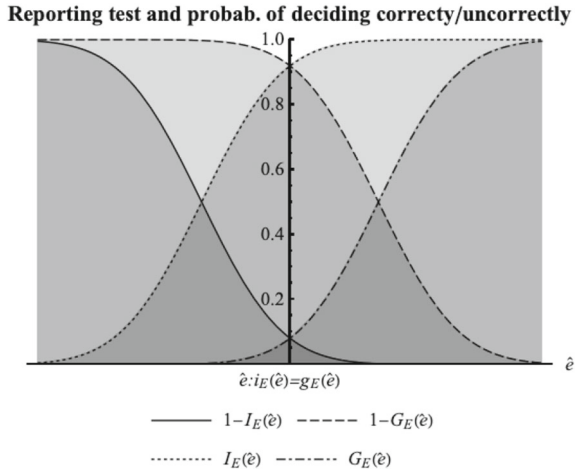
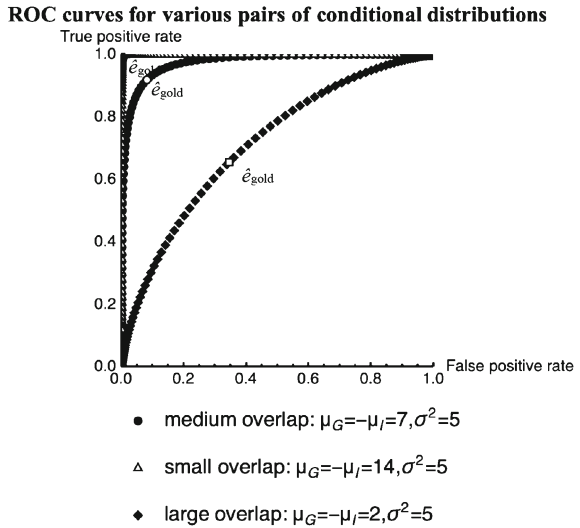


Fig. 2 ROC curve



The assumption of symmetry with respect to a vertical axis remains reasonable because we presume that the process of evidence production is analogous both in the case of guilty behavior and in the case of legitimate transactions. However, clients' behavior could differently affect the availability of indications of innocence or guilt contingent on whether they are innocent or guilty. Elusive behavior on the side of guilty clients might imply conditional distributions that are not symmetric with respect to a vertical axis. The proofs in Appendix 1 are provided by assuming two generic distributions that are symmetric with respect to any given vertical axis (in the figures provided in the rest of the paper, the conditional distributions are symmetric with respect to the vertical axis $e = 0$). By inspecting the proofs provided in Appendix A.1.1 we observe that symmetry is needed only to provide clearer and intuitive thresholds

Fig. 3 Reporting test and money-laundering rate for different levels of sanctions

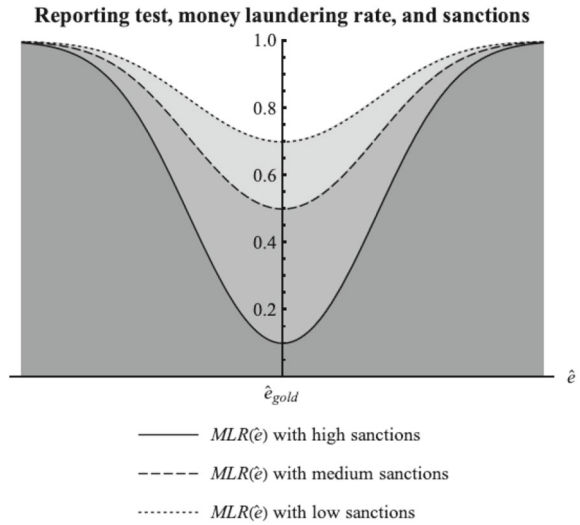
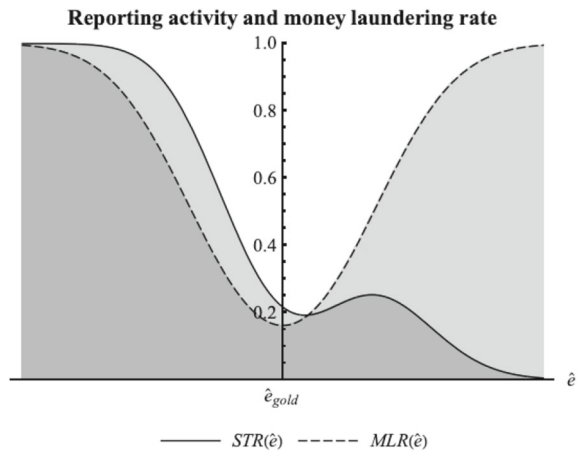


Fig. 4 Reporting activity and money-laundering rate



and the results of Sect. 2.4. Implications for the gold standard and deterrence (Result 1) and the relation between error rate, reporting test, and sanctions (Result 2) are valid even by relaxing symmetry.

2.2 The Reporting Test and the Decision to Issue an STR

For each transaction, financial intermediaries and professionals observe a certain net evidence e and issue an STR when e is greater than a certain threshold, \hat{e} , which is the reporting test. Therefore, *given a transaction*, the probability of a correct STR being issued is $P[E > \hat{e} | \textit{Guilty}] = 1 - G_E(\hat{e})$. Conversely, an STR is incorrectly issued (type-I error) with a probability of $P[E > \hat{e} | \textit{Innocent}] = 1 - I_E(\hat{e})$. Moreover, an

Fig. 5 Error rate, shares of type-I and type-II errors, and money-laundering rate

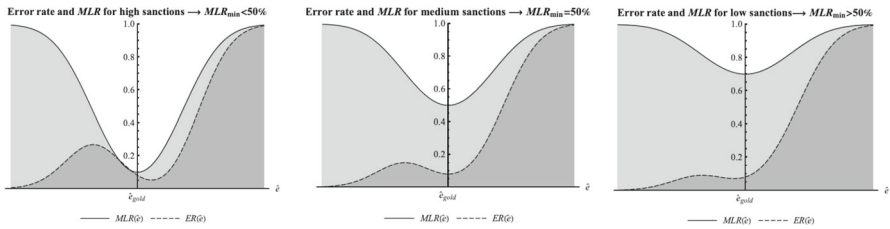
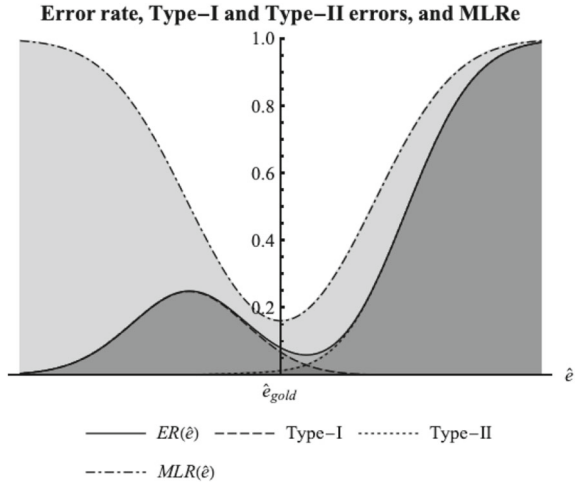


Fig. 6 Error rate for various MLR_{min} corresponding to various sanctions S

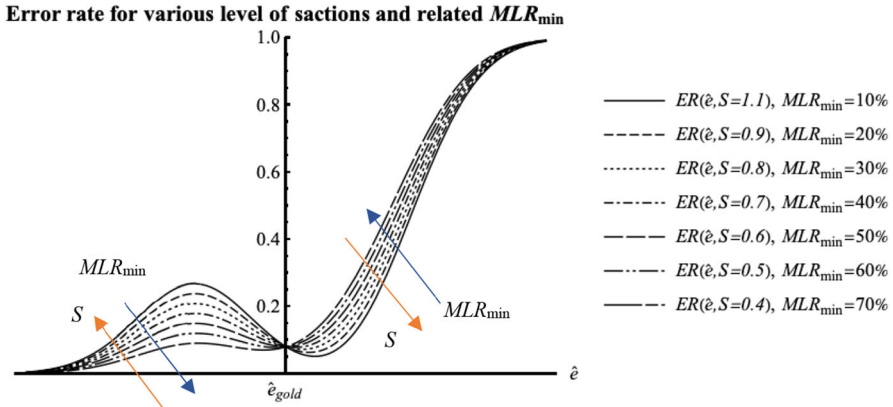


Fig. 7 Error rate for various MLR_{min} corresponding to various sanctions S

Fig. 8 Intervals of assumptions
H1–H3

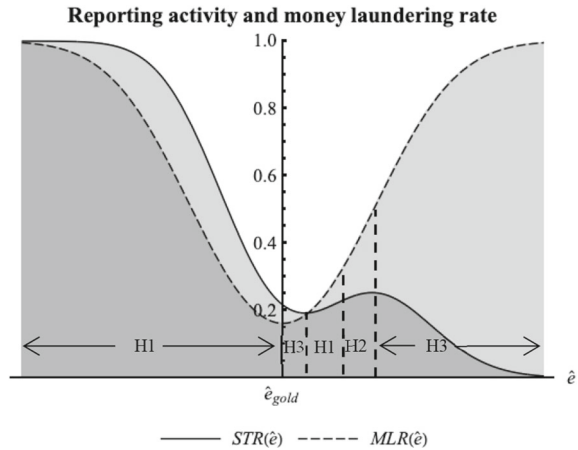
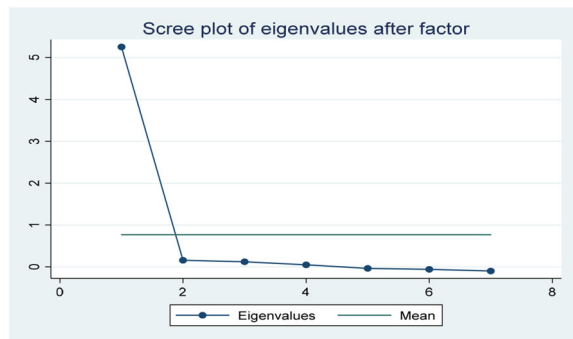


Fig. 9 Scree plot from factor analysis on predicate crimes, Italian provinces, 2009–2012



STR is correctly not issued for a transaction that is not an act of money laundering with a probability of $I_E(\hat{e})$. Finally, an STR is incorrectly not issued for a money-laundering transaction (type-II error) with a probability of $G_E(\hat{e})$. Figure 1 illustrates these probabilities as functions of the reporting test \hat{e} .⁵

Note that the probabilities of committing a type-I or a type-II error for a given transaction depend on the overlap between $i_E(e)$ and $g_E(e)$. As shown by the ROC curves (Marzban 2004) plotted for different pairs of conditional distributions in Fig. 2, the usefulness of the reporting test (as measured by the area under the ROC curve) depends on the extent to which the conditional distributions overlap and, in turn, how the evidence is either confounding or discriminating. The *gold standard* reporting test that maximizes the probability of reporting a true positive while simultaneously minimizing the probability of reporting a false positive⁶ is such that the conditional probability

⁵ Probabilities in Fig. 1 have been plotted by assuming the following parameters for the Normal conditional distributions of E : $\mu_G = -\mu_I = 7$ and $\sigma_I^2 = \sigma_G^2 = 5$.

⁶ See Zou et al. (2007): an ROC curve is a plot of sensitivity (true positive rate) on the y axis against $(1 - \text{specificity, false positive rate})$ on the x axis for varying values of the reporting test. The 45° diagonal line connecting (0,0) to (1,1) is the ROC curve corresponding to random chance. The ROC curve for the gold

distributions i_E and g_E cross each other, that is $\widehat{e}_{gold} : i_E(\widehat{e}_{gold}) = g_E(\widehat{e}_{gold})$. A formal proof is provided in Appendix 1.⁷

2.3 Reporting Test, Sanctions, and Money Laundering Activity

The possibility of being caught when laundering money represents a deterrent to money laundering activities. Deterrence depends on several factors, particularly the effectiveness of the AML system and imposed sanctions. As we focus on the first level of the AML system, we model the deterrent impact of a given reporting test \widehat{e} by considering the related probabilities to be reported as a suspicious client, *ceteris paribus*; that is, we do not explicit other factors, including the probability of being investigated by the police as a consequence of independent investigations or the probability that the further levels of the AML system decide to discharge intermediaries' reports.⁸

We define the stochastic variable $w > 0$ as the extra gain that a money laundering transaction would produce, net of the gain that could otherwise be obtained legally (i.e., by obeying the law). Then, we define $S > 0$ as the expected sanction applied when an STR is issued.⁹

A rational, risk-neutral individual undertakes a money-laundering transaction only when the associated expected benefit is greater than what is expected from abstaining and being eventually erroneously sanctioned. Thus, a money laundering transaction is committed when

$$\begin{aligned}
 w - S[1 - G_E(\widehat{e})] &> -S[1 - I_E(\widehat{e})] \\
 w &> S[I_E(\widehat{e}) - G_E(\widehat{e})] \equiv \widehat{w}(S, \widehat{e})
 \end{aligned}
 \tag{1}$$

Given (1), the probability that an individual decides to engage in a money laundering transaction is

$$P[w > \widehat{w}(S, \widehat{e})]
 \tag{2}$$

As we do not know the probability distribution of w , we restrict ourselves to defining a probability measure¹⁰ consistent with the probability defined in (2). In particular, starting from the deterrence threshold $\widehat{w}(S, \widehat{e})$ expressed in (1), we define a probability

Footnote 6 continued

standard is the line connecting (0,0) to (0,1) and (0,1) to (1,1). The area under the ROC curve is a summary measure that essentially averages diagnostic accuracy across the spectrum of tests.

⁷ Since in the simulations that we used to produce the figures of the paper we assume that the conditional distributions cross at 0, here the gold standard reporting test is $\widehat{e}_{gold} = 0$.

⁸ Nevertheless, we account for the probability of being caught in the money laundering activity through the inclusion of the variable "Detected ML crimes/pop" (ML crimes with known offender /total ML crimes) as a covariate in the regressions. Furthermore, we will also account for the fact that lengthy procedures and delays in trials dilute the incidence of both sanctions and crime deterrence.

⁹ In the real world, sanctions for money laundering are applied by the third level of the AML system, which is the judicial system. The expected sanction S under consideration is an estimation of the sanction applied by the judicial system conditional on the issuance of an STR.

¹⁰ A *probability measure* on a set is a systematic way to assign a number to each suitable subset of that set, which is intuitively interpreted as its size. It takes the value of 1 on the whole space (and, therefore, takes all its values in the unit interval [0, 1]).

measure such that the condition determining the probability expressed in (2) is verified for a given value of w . Specifically, we define the *money-laundering rate*¹¹ (MLR) as

$$MLR(S, \hat{e}) = 1 - \hat{w}(\hat{e}) = 1 - S[I_E(\hat{e}) - G_E(\hat{e})] \quad (3)$$

As expected, the money laundering rate negatively depends on sanctions S (see Fig. 3 and proof A.1.2. in Appendix 1). As illustrated in Fig. 3, for any level of sanctions S , the money laundering rate $MLR(\hat{e})$ is minimal when the reporting test is set equal to the *gold standard* \hat{e}_{gold} (see proof A.1.3 in Appendix 1).¹²

Result 1 (Deterrence) The reporting test that minimizes the money-laundering rate is the gold standard $\hat{e}_{gold} : i_E(\hat{e}_{gold}) = g_E(\hat{e}_{gold})$.¹³ Sanctions negatively affect the money laundering rate.

2.4 Reporting Activity, Errors, and Accuracy

As shown in Sect. 2.3, by setting the reporting test, signaling bodies can affect the money laundering rate. If they opt for the gold standard, they maximize deterrence. However, the reporting test implemented by intermediaries depends on their aims. Different elements concur in the formation of the reporting test used by the signaling bodies: primarily, the bank's incentive to monitor and report and the actual possibility to identify the true nature of a transaction (Takàts 2011). Monitoring and reporting are costly tasks for signaling bodies. Similarly, mistakes in these dual tasks are costly because type-II errors are typically sanctioned by the upper levels of the AML system, and type-I errors can imply relevant reputational costs. Although it is not obvious that type-I and type-II errors should be equally weighted from the intermediaries' perspective and even from a social welfare maximization perspective, given the overall approach of the AML system described in Sect. 1, we assume that the pivotal objective for signaling bodies is to maximize the accuracy of their reporting activity.

Given the sanctions S and the reporting test \hat{e} , money laundering occurs with probability $MLR(S, \hat{e})$. This implies that financial intermediaries and professionals that observe all the transactions that take place report a share of transactions to the FIU as suspicious, corresponding to

$$STR(S, \hat{e}) = \underbrace{MLR(S, \hat{e}) \times (1 - G_E(\hat{e}))}_{\text{Share of correct reports}} + \underbrace{(1 - MLR(S, \hat{e})) \times (1 - I_E(\hat{e}))}_{\text{Share of type-I errors}} \quad (4)$$

where the first addendum corresponds to the share of truly positive reports, and the second addendum corresponds to the share of falsely positive reports (type-I errors).

A further relevant measure for the AML system is the error rate $ER(\hat{e})$, which is defined as the share of total transactions that are either erroneously reported as suspicious (type-I errors) or erroneously not reported though implying money laundering

¹¹ To guarantee that the MLR is between 0 and 1, we set $0 < S < 1/[I_E(0) - G_E(0)]$.

¹² Figure 3 is based on the usual assumptions about the conditional probability distributions of the net evidence; sanctions S are set equal to 1.1, 0.6, and 0.4.

¹³ Here, $\hat{e}_{gold} = 0$.

(type-II errors) by the signaling bodies:

$$ER(\hat{e}) = \underbrace{(1 - MLR(\hat{e})) \times (1 - I_E(\hat{e}))}_{\text{Share of type-I errors}} + \underbrace{MLR(\hat{e}) \times G_E(\hat{e})}_{\text{Share of type-II errors}}$$

$$= S[I_E(\hat{e}) - G_E(\hat{e})](1 - I_E(\hat{e})) + (1 - S[I_E(\hat{e}) - G_E(\hat{e})])G_E(\hat{e}) \quad (5)$$

We study $STR(S, \hat{e})$ and $ER(S, \hat{e})$ as functions of the reporting test \hat{e} (see Fig. 4) and sanctions S . On this concern, note that banks and intermediaries (the first level of the AML system) can adjust the reporting test that they employ in their AML activities while considering sanctions S as exogenous.

Regarding the reporting activity STR , it is decreasing in sanctions S since higher sanctions discourage money laundering (see Appendix A.1.4). Instead, STR is not a monotonic function of the reporting test; it is initially decreasing and then locally increasing when the reporting test is sufficiently higher than the gold standard to induce intense money-laundering activity that, in turn, pushes reporting. Finally, when the reporting test increases even further, reports decrease, although money laundering is intense because the strictness of the test prevails.

It is worth noting that our analysis can be limited to the neighborhood of the gold standard \hat{e}_{gold} . On the one hand, reporting tests in the neighborhood of the gold standard are sufficiently accurate for each given transaction (see the ROC curve in Fig. 2) and guarantee a good level of deterrence (see Fig. 3). On the other hand, as showed in Fig. 4, "extreme" reporting tests (either very low or very high with respect to the gold standard) result in limit cases: a very low reporting test ($\hat{e} \ll \hat{e}_{gold}$) implies that all the transactions are correctly reported as suspicious (no errors) because any incentive to abstain from money laundering completely disappears (money laundering activity explodes, as shown in Fig. 3). Analogously, a very high reporting test $\hat{e} \gg \hat{e}_{gold}$ implies that, although all the transactions are illicit, none is reported (100% of errors).¹⁴ The proofs are provided in Appendix A.1.4.

Concerning the errors that signaling bodies make, we observe that $ER(\hat{e})$ is decreasing in sanctions S when the applied reporting test is smaller than the gold standard and increasing in sanctions S when the applied reporting test is greater than the gold standard (see Appendix A.1.5).

$ER(\hat{e})$ is the sum of the share of type-I errors and the share of type-II errors (Fig. 5). By inspecting (5), we observe that when the reporting test is low, the share of type-I errors prevails (the second addendum is initially null and then very small because $G_E(\hat{e} \ll \hat{e}_{gold}) \rightarrow 0$); conversely, when the reporting test is high the share of type-II errors prevails (the first addendum is very small and finally null because $1 - I_E(\hat{e} \gg \hat{e}_{gold}) \rightarrow 0$). This is consistent with the general idea that stricter tests imply many false negative reports, whereas excessively lax tests result in many false positive reports.

As illustrated in Fig. 5 (and proven in Appendix A.1.5), $ER(\hat{e})$ shows a global maximum and a global minimum. Both these points do not deserve particular comments as they correspond to reporting tests either so low or so high with respect to the gold standard that they result in the limit cases described above: $\hat{e} \ll \hat{e}_{gold}$ implies that all

¹⁴ In Fig. 4 sanctions S are set equal to 1, the other parameters are the usual ones.

the transactions are correctly reported (0% of errors) because any incentive to abstain from money laundering completely disappeared; $\hat{e} \gg \hat{e}_{gold}$ implies that, though all the transactions are illicit, none is reported (100% of errors).

Focusing on the neighborhood of the gold standard \hat{e}_{gold} and defining \hat{e}_{minER} as the reporting test that *locally* minimizes the error rate, we prove that $ER(S, \hat{e})$ has its local minimum in \hat{e}_{gold} if and only if $S = \frac{0.5}{(I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold}))}$ (see Appendix A.1.5). In addition, we verify that when $S < \frac{0.5}{(I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold}))}$ the local minimum of the error rate is on the left of \hat{e}_{gold} ; the smaller the sanctions, the lower the local minimum of the error rate. Conversely, when $S > \frac{0.5}{(I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold}))}$ the local minimum of the error rate is on the right of \hat{e}_{gold} ; the bigger the sanctions, the higher the local minimum of the error rate.

Finally, recalling Eq. (3) and result 1, we define $MLR_{min} \equiv 1 - S[I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold})]$. Now, we can rewrite the relation between sanctions and MLR_{min} as $S = \frac{1 - MLR_{min}}{I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold})}$. This allows us to reformulate the previous implications as follows:

$$\begin{aligned} \text{if } MLR_{min} > 0.5, \hat{e}_{minER} &< \hat{e}_{gold} \\ \text{if } MLR_{min} = 0.5, \hat{e}_{minER} &= \hat{e}_{gold} \\ \text{if } MLR_{min} < 0.5, \hat{e}_{minER} &> \hat{e}_{gold} \end{aligned} \quad (6)$$

Moreover, by following the same reasoning made for sanctions, we remark that the higher the $MLR_{min} > 0.5$, the lower the local minimum of the error rate; the lower the $MLR_{min} < 0.5$, the higher the local minimum of the error rate. Figure 6 provides examples of the usual parameters.

Finally, the analysis above allows to derive the following:

Result 2 (Accuracy of reporting activity) The reporting test \hat{e}_{minER} that (locally) minimizes the error rate decreases as the minimum money-laundering rate MLR_{min} increases (sanctions S decrease). $\hat{e}_{minER} = \hat{e}_{gold}$ only for a specific level of MLR_{min} (sanctions S).

As expressed, result 2 is quite general, while the relevant threshold for $MLR_{min} = 50\%$ (corresponding to the sanction threshold $S = \frac{0.5}{(I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold}))}$) strictly depends on the assumptions of symmetric conditional distributions of g and i . Figure 7 graphically summarizes the present implications by assuming the usual parameters.

As illustrated by the model, intermediaries and professionals typically face a trade-off when setting their reporting tests. While deterrence goals can always be pursued by setting the reporting test to the gold standard $\hat{e} : i_E(\hat{e}) = g_E(\hat{e})$, the accuracy of reporting activity depends on the minimum level of money-laundering activity (which in turn depends on sanctions). Those intermediaries and professionals concerned with the accuracy of their reporting activity should opt for a reporting test lower than the gold standard when sanctions are low and money-laundering activity is intense. Conversely, they should employ a reporting test higher than the gold standard when sanctions are high and money-laundering activities are not overly pervasive.

3 Empirical Analysis

The empirical analysis aims to assess the evolution of the quality of the STR received from financial intermediaries and processed by the Italian FIU to understand whether the risk-based approach introduction has changed the reporting test adopted by intermediaries at the first level of the AML system.

Two considerations are worthwhile before conducting the empirical exercise. First, the analysis focuses on the role of financial intermediaries in combating the phenomenon of money laundering. Banks and other financial companies play a predominant role in the transmission of suspicious transactions. For example, in the period on which our analysis focuses (2009–2012) financial intermediaries transmitted 48,822 STRs to the Italian FIU, against 805 STRs sent by professionals and non-financial operators.¹⁵ No STR was received from other potential reporting agents. Furthermore, given the size and quantity of transactions processed, financial intermediaries have developed sophisticated technologies and software to improve the quality of reporting activity. These instruments are infrequently used by other reporting agents, in light of the economies of scale required to use them.

Second, by quality of STR, we refer to the possibility of financial intermediaries to (i) adjust the share of errors they commit, and (ii) affect ML deterrence in the design of the suspicious reporting mechanism.

Through the predictions provided by the model illustrated in the previous section, the empirical analysis is primarily aimed at understanding how possible changes in the reporting activity due to the introduction of the risk-based approach and KYC principles may have affected the number of STRs (for a given volume of banking activities), the money laundering rate (identified as the dimension of laundering activity in relation to predicate crimes), and consequently type-I error (false positives reported by the banks as suspicious transactions to the FIU). It is therefore important to note that this study is not aimed to estimate the volume of ML, but rather focuses on the evolution of the behavior of intermediaries in the process of reporting STRs to the FIU.

To empirically evaluate this aspect, we perform regression analysis on Italian data, referring to the period immediately following the introduction of the risk-based approach (which took place in 2009 in Italy) at the provincial level. This will also allow to account for geographical specificities that might characterize the incidence of both laundering and reporting activities (dalla Pellegrina et al. 2020a, b). We use regression analysis, based on the use of proxies of the reporting test, STR, and money laundering rate (MLR) along with the model's predictions to orient the statistical inference regarding the Italian situation in the period after 2009. Finally, by combining the empirical evidence obtained through regression analysis with the time evolution of aggregated data (at the national level), we will provide inference in terms of type-I error, and from there, we will use the model to infer the pattern of type-II error in the period of interest.¹⁶

¹⁵ Data were provided by the Italian FIU.

¹⁶ In this case, given the absence of useful data to formulate a proxy for the type-II error, the model is the only tool available to evaluate the effects of the introduction of the risk-based approach.

3.1 Empirical Hypotheses

In the first part of the analysis, we check the relationship between *i*) a proxy for the reporting test and the actual evolution of STR in terms of the volume of bank activity, and *ii*) a proxy for the reporting test and the MLR.

Referring to the model in Sect. 2, we make the following assumptions:

H1. If both (i) and (ii) indicate a negative correlation between the variables of interest, we infer that the reporting test is lower or in the positive neighbors of the *gold standard* (see interval H1 in Fig. 8).

H2. If both (i) and (ii) indicate a positive correlation between the variables of interest, we infer that the reporting test is higher than the gold standard (interval H2 in Fig. 8).

H3. If (i) indicates a negative correlation between STR and the reporting test, while (ii) indicates a positive correlation between MLR and the reporting test, we can infer that the latter is much higher than the gold standard (or in a narrow interval at the immediate right-hand side of the gold standard) (intervals H3 in Fig. 8).

However, the assumptions outlined above do not represent sufficient conditions to understand the evolution of the variables of interest, that is, STR and MLR in the period following the introduction of the risk-based approach. In fact, referring again to Fig. 8, we are not aware of the increasing or decreasing pattern of the reporting test after 2009, as we do not know if the intermediaries have decided to raise or lower the reporting test.

To better understand these aspects, the official statistics reporting trends over time of (our proxies of) the reporting test, STR, and MLR are of fundamental support. In particular, if both STR and MLR increase over time and at the same time, the proxy of the reporting test decreases, we can infer that H1 is verified. The same will be true if both STR and MLR decrease over time, and at the same time, the proxy of the reporting test is increasing.

On the contrary, if both STR and MLR increase over time and at the same time, the proxy of the reporting test increases, and we can infer that H2 is verified. The same will be true if both STR and MLR decrease over time, and at the same time, the proxy of the reporting test decreases. If none of these cases occur, hypothesis H3 holds.

According to the arguments presented in Sect. 1, we expect that H1 may plausibly hold; in particular, that financial intermediaries have relaxed the reporting test as a consequence of the imbalance between pro and cons of the reporting activity by intermediaries (“crying wolf” attitude). This would mean that banks now ask for less evidence to report suspicious transactions to the FIU, and consequently, the number of bad reports may have reasonably increased. Using official data from the FIU, we check whether STRs to the FIU have actually increased over time in relation to the banking activity, and especially if this increase in reports has translated into an increased rate of dismissal of reports received (at an aggregate level), a figure officially recognized as a measure of type-I error.¹⁷ From here, we proceed with the appropriate inference of the type-II error using only the predictions of the model.

¹⁷ The more reliable measure used to measure type-I error is represented by the STR dismissal rate (FIU Annual Reports 2010–2013).

3.2 Data Sources and Proxies for Reporting Test and STR

The data collected from the Italian provinces cover the period from 2009 (corresponding to the introduction of the risk-based approach in Italy) to 2012.¹⁸

A challenging methodological aspect stems from the fact that we cannot directly observe the true incidence of money laundering because of its concealed nature. We use the number of police reports for ML as a proxy of the amount of assets laundered by criminals. As widely debated in the literature, this measure has to be taken cautiously as criminal activities are part of an underground economy, thus the number of reports submitted to the authorities provides only partial insight into the overall phenomenon. Nevertheless, the total number of money laundering police reports is generally recognized to be a good indicator of the actual (unobserved) flow of laundered money. The underlying hypothesis is that all other factors being equal (i.e. controlled for in empirical analyses, with particular emphasis on predicate crimes), there should be positive correlation between police reports and the actual incidence of ML (see dalla Pellegrina et al. 2020a, b, and literature cited therein).¹⁹

The number of suspicious transactions was provided by the FIU at the provincial level for the period of interest. We were endowed with STRs from financial intermediaries, representing the most consistent volume of reports transmitted to the FIU.²⁰ Data are bi-annual, and we summed the number of reports accrued to the FIU in each semester. We used the lagged value of the number of reports made in the last 12 months in relation to each predicate crime to build an index representing a proxy for the reporting test by means of factor analysis.

As being part of the criminals' utility function, the probability to get caught when laundering money is accounted for by including a proxy of the detected ML crimes, divided by each province's population (ML crimes with known offender /total ML crimes). Furthermore, as is known in the literature, lengthy procedures and delays in trials dilute crime deterrence. We used the duration of the criminal trials (days) to capture the deterrence effects against ML and predicate crimes. Data were obtained from the Ministry of Justice on an annual basis and disaggregated by district and court. It was therefore aggregated to have an average duration of the processes by district and reworked to have a correspondence between district and province.

We also account for the fact that, for reasons of non-traceability, one of the main mechanisms through which ML is executed involves cash transactions (Axelrod 2017;

¹⁸ We concentrate on the years immediately following the introduction of the risk-based approach. We have more recent data from the FIU, but this information is aggregated at the national level. Nevertheless, our aim is to observe the effects of the introduction of the risk-based approach following the implementation of the first EU directive on this matter (Directive 2005/60/EC on the prevention of the use of the financial system for the purpose of money laundering and terrorist financing, Legislative Decree No. 231 of November 21). For this reason, we believe it is crucial to concentrate on the years immediately after the introduction of the new legislation, as the adjustment process implemented by financial intermediaries and other reporting agents, and especially changes in the screening procedures adopted by the FIU (see end of Sect. 3.3) no longer require an extended time period and produce statistically significant effects. In addition, amendments of the legislation have been introduced, which could interfere with the effects investigated in this study.

¹⁹ We are aware of the fact that a work more focused on estimating the incidence of money laundering should use more sophisticated methodologies than ours. See, for instance, Ferwerda et al. (2020).

²⁰ See beginning of Sect. 3.

Dalla Pellegrina and Masciandaro 2009; Lowe 2017). We include cash inflows to the banking system (finally accruing to the Central Bank) divided by each province's population. In the same spirit, as they could favor cash transactions compared to other forms of payment (for example, online), we also add the number of bank branches present in each province. In line with the above, both the number of real estate transactions and the number of real estate loans were included in the regression analysis (all data are from the Bank of Italy on an annual basis). Real GDP and real per capita GDP are also included to account for the (economic) dimension of each province (ISTAT). The summary statistics for all variables involved in the analysis are presented in Appendix 2.

To construct a measure of reporting tests, we collected information on police reports of both money laundering and predicate crimes for each Italian province on an annual basis. Previous studies (e.g., Abadinsky 2010) estimate that organized crime's highest-return activities are drug trafficking, exploitation and abetting of prostitution, racketeering, fraud, and counterfeiting of brands and industrial products. In line with the literature (e.g., Arnone and Borlini 2010; Draghi 2007; Jayasekara 2020; Mugarura 2011), we added armed robbery and micro-crime indexes. In addition, to capture the activities conducted by criminal organizations, we also included criminal associations, mafia-type associations, usury, and corruption. Finally, the tax gap is used as a measure of tax evasion. All data were obtained from the Italian Institute of Statistics (ISTAT). In particular, police reports on various crimes were drawn from criminal judicial statistics, while the tax gap was extracted from the well-being and sustainability indicators (BES).

We define the following empirical proxy for the reporting test \hat{e} :

$$\hat{e} = -factor\left(\frac{STR}{sourcecrime_k}\right) \quad (7)$$

The numerator accounts for excess reporting activity to the FIU in a given province-year. Reporting activity was measured as the number of STRs submitted to the FIU by financial intermediaries. Predicate crimes are used in the denominator to normalize the excess reporting activity across provinces with different crime rates, and k refers to each predicate crime illustrated above.

As there are several different types of predicate crimes, we performed a confirmatory factor analysis (Jöreskog 1969) to obtain a unique and comprehensive measure of the standard of error. This technique is useful to the extent that the frequencies of similar types of crime are correlated across provinces. Hence, reducing the number of predicate crimes to one or more latent factors simplifies the interpretation of the subsequent empirical analysis.²¹ Specifically, we constructed several individual reporting test measures, one for each type of predicate crime. These measures have identical numerators, but different denominators. The factor analysis of these individual measures allows us to obtain a reduced number of factors as proxies for the reporting test.

²¹ Alternatively, we could have summed the number of police reports for the different types of crimes to construct a synthetic measure of predicate-crime activity. However, factor analysis helps give more weight to those predicate crimes that play a greater role in the money-laundering process.

Table 1 Factor analysis on predicate crimes, Italian provinces, 2009–2012

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor 1	5.18	3.66	0.76	0.76
Factor 2	1.52	1.26	0.22	0.99
Factor 3	0.26	0.08	0.04	1.03
Factor 4	0.18	0.03	0.03	1.05
Factor 5	0.14	0.10	0.02	1.07
Factor 6	0.04	0.01	0.01	1.08
Factor 7	0.03	0.05	0.00	1.09
LR test: independent versus saturated: $\chi^2(21) = 3287.15$, Prob. > $\chi^2 = 0.00$				
Number of observations	412			
Retained factors	7 (only retained factors are displayed)			
Number of parameters	70			

Note that the retrieved factor in Eq. (7) is an inverse measure of what is referred to in the model in Sect. 2 as the reporting test ($\hat{\rho}$). To obtain a measure of the reporting test, which is in line with the model’s meaning, we take the negative value of the factor as a proxy of $\hat{\rho}$. Specifically, its increase identifies the fact that banks need more evidence to report a transaction as suspicious to the FIU.

The results of the factor analysis are shown in Table 1 (additional details are provided in Appendix 3). The corresponding scree plot presented in Fig. 9 suggests the retention of a single factor that represents a unique empirical measure of reporting tests at the provincial level. We aim to analyze its evolution over time and its incidence on other measures in the model’s setup.

3.3 Econometric Evidence

According to Sect. 3.1 and with specific reference to the hypotheses outlined therein, the first step in assessing the model’s ability to evaluate the effects of changes in the reporting test on STR, MLR, type-I, and type-II errors is to understand where in Fig. 4 the data suggest setting the Italian situation in the years of interest. More precisely, we aim to formerly check the sign of the correlation between the reporting test illustrated in ($\hat{\rho}$) and STR, on the one hand, and between $\hat{\rho}$ and the MLR, on the other hand, at the provincial level.

To this end, we estimate the following regressions:

$$STR_{it} = \alpha_0 + \alpha_2 \hat{\rho}_{it} + \alpha_3 RealGDP_{i,t} + \alpha_4 RealpercapitaGDP_{i,t} + \alpha_5 Lengthofcriminaltrials_{i,t} + \alpha_6 Detected_ML_{i,t} + \lambda_i + \mu_t + \epsilon_{it}, \quad (8)$$

$$MLR_{it} = \beta_0 + \beta_2 \hat{\rho}_{it} + \beta_3 RealGDP_{i,t} + \beta_4 RealpercapitaGDP_{i,t} + \beta_5 Lengthofcriminaltrials_{i,t} + \beta_6 Detected_ML_{i,t} + \lambda_i + \mu_t + \epsilon_{it}, \quad (9)$$

where the subscript i refers to the province and t refers to the year.

Equations (8) and (9) are the baseline specifications. We provide additional regressions by adding cash transactions, bank branches, and real estate-related variables. Other specifications involve a set of predicate crimes divided by each province's population.

Given our empirical purposes, we measure STR as the number of suspicious transactions reported by banks to the FIU scaled by the volume of banks' activities, namely the amount of bank transactions (loans plus deposits). MLR is the money laundering rate observed in the province in the 12 months preceding time t , measured by the number of police reports for ML in relation to the economic dimension of each province (GDP)²²; λ_i and μ_t are province and time fixed effects, respectively and $\varepsilon_{i,t}$ is an idiosyncratic error term clustered at the provincial level. The estimated signs of α_2 and β_2 are our focus, as they capture the correlation between MLR and our proxy for the reporting test to address our hypotheses.²³

The regression output, which is obtained through linear estimation, is reported in Table 2 (upper panel). The negative sign of the proxy of the parameter associated to $\hat{\varepsilon}$ indicates that either as $\hat{\varepsilon}_{i,t}$ increases both STR and MLR decrease, or as $\hat{\varepsilon}_{i,t}$ decreases both STR and MLR increase. This is likely supportive of H1, suggesting that in Italy $\hat{\varepsilon}$ is likely to be positioned on the left-hand side or in the neighbours of the gold standard (i.e., not far from the golden rule in Fig. 8).

To elucidate each possible reverse situation (i.e., whether the reporting test has actually increased or decreased in the observed period), we consider the pattern of the aggregate variables at the national level. In other words, we use the model's predictions combined with the statistics in Table 3 to make inferences on the direction taken by the reporting test in Fig. 8.

First, Table 3 (Column (1)) shows that our proxy for $\hat{\varepsilon}$ followed a substantially decreasing pattern from 2009 to 2012, indicating that the reporting test decreased in the aftermath of the introduction of the risk-based approach.

Second, our data point toward a growing number of STRs in relation to the volume of bank transactions (STR) in the observed period (Table 3, Column (2)). Since STRs are not monotonous for all positive values of the reporting test, an increase in reporting activity combined with a decrease $\hat{\varepsilon}$ confirms that H1 may reasonably hold.

Third, Table 3 (Column (3)) provides evidence of the stability of the MLR . According to the model, the MLR was around its minimum in the period of interest, that is, $\hat{\varepsilon}$

²² The use of GDP as denominator in the empirical measure of the MLR, partially accounts for the possibility that money is not laundered locally. Typically, criminal infiltration in the legal economy is computed as the share of revenues from source crimes transformed into legal money, hence source crimes are used to scale the number of ML crimes (see dalla Pellegrina et al. 2020a, b) to measure the vulnerability to ML of a geographical area in regression analyses. However, this implicitly assumes that money can only be laundered locally, that is, within the area's borders. To overcome this constraint, we use GDP as a scale-measure to account for the possibility of conducting money laundering abroad, i.e. export laundering activities to other provinces. Indeed, for given aggregate local income (at the area's level), GDP accounts for net export. The assumption is that the larger net export of goods and services, the higher the probability to export (also) dirty money to be laundered.

²³ The regression analysis is not intended to estimate a causal relationship between the reporting test and the money-laundering rate. The sign of the correlation between the two variables only aims to provide an understanding of Italy's position in Fig. 8.

Table 2 Regression analysis, Italian provinces, 2009–2012

	Dependent variable					
	(1)	(2)	(3)	(4)	(5)	(6)
	STR	STR	STR	MLR	MLR	MLR
$\hat{\epsilon} = \text{-factor 1}$	- 0.044*** (0.003)	- 0.043*** (0.003)	- 0.049*** (0.001)	0.019* (0.007)	0.015 (0.008)	0.019** (0.007)
Duration of criminal trials	0.008 (0.009)	0.005 (0.009)	0.011 (0.006)	- 0.025 (0.057)	- 0.030 (0.057)	0.036 (0.063)
Detected ML crimes/pop	- 0.001 (0.001)	- 0.000 (0.001)	- 0.001 (0.001)	- 0.014 (0.015)	- 0.012 (0.016)	- 0.012 (0.013)
Per capita GDP	- 1.922 (2.019)			- 33.38* (14.670)		
GDP	0.114 (0.160)			2.707* (1.201)		
Bank branches/pop		- 0.095 (0.610)			- 1.994 (4.885)	
Cash inflows/pop		0.183 (0.140)			- 0.328 (0.969)	
Real estate transactions/pop		- 0.134 (0.179)			- 0.036 (1.180)	
Real estate loans/pop		- 0.092 (0.122)			0.136 (0.753)	
Constant	0.524 (0.272)	0.370 (0.314)	- 0.003 (0.061)	- 2.191 (2.028)	1.485 (2.369)	- 0.181 (0.720)
Predicate crimes /pop ^(a)	No	No	Yes	No	No	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.98	0.98	0.99	0.73	0.72	0.77
N	412	412	412	412	412	412

Linear regression model. Standard errors clustered at provincial levels in parentheses. All variables are in log form. See Appendix 2 for details on individual predicate crimes

*p < 0.10, **p < 0.05, ***p < 0.01

^aSuspicious Transactions and Reporting Test

was likely to be closer to the gold standard. More precisely, from the pattern of MLR, which decreased from 2009 to 2010, stabilized in 2011, and then increased in 2012, we can conclude that deterrence was most likely maximized in 2010–2011, when it crossed the gold standard (from right to left, see Fig. 8).

Table 3 Patterns of \hat{e} , MLR and STR, Italian provinces, 2009–2012

Year	Reporting test (\hat{e}) ⁽¹⁾ pattern in the period of interest (1)	STR ⁽²⁾ pattern in the period of interest (2)	MLR ⁽³⁾ pattern in the period of interest (3)
2009	0.799	0.258	0.120
2010	0.166	0.291	0.118
2011	− 0.308	0.319	0.118
2012	− 0.657	0.338	0.124

(1) Equal to *-factor* 1 (averaged by year)

(2) Measured as total STRs submitted to the FIU/volume of bank transactions (at the national level)

(3) Measured as ML police reports/volume of bank transactions (at the national level)

All data are in logs (Log (1 + var))

To provide inference on type-I and type-II errors, we use the sufficient statistics approach pioneered by Chetty (2009) and recently applied to the study of money laundering by Imanpour et al. (2019). One useful feature of the model is that the variables of interest depend on only a few constructs that correspond to the real world and are easily observable in the available (mostly aggregated) data.

To test the predictions of the model, we refer to Fig. 5 in Sect. 2, suggesting that type-I error should have increased in the aftermath of the introduction of the risk-based approach, which we motivated as a consequence of the reduction of the reporting test (i.e., a shift leftwards in Figs. 5 and 8, in the neighbours of the gold standard).

Some (aggregate, at national level) measures of reports received, processed and dismissed by the FIU are available from the FIU Anti-money laundering notebooks (Statistical data and annual reports). Precisely, we have been endowed with three types of reports: (i) total reports for suspicious transactions (money laundering, terrorism, arms trafficking) received from financial intermediaries; (ii) reports for ML received from all reporting agents (professionals and non-financial operators, others), and (iii) dismissed reports for money laundering received from all reporting subjects, including professionals and non-financial operators. We are not endowed with a precise measures of the reports for ML received from financial intermediaries, but just with the total number of STR, including those for terrorism and arms trafficking, although the latter represent only 1% of the STR from intermediaries. Hence we believe this can be interpreted as a reliable measure of the flow of reports for ML received from financial intermediaries.

Starting from these data, we constructed some different proxies of type-I error, all being representative measures of the dismissal rate of the STR received by the FIU, with focus on financial intermediaries. As reported in Table 4, the ratio of reports received but not analyzed by the FIU to the overall reports received from intermediaries (column (1)) increased from 2009 to 2011 and then returned to 2008 levels. Similarly, the overall dismissed reports for ML to the analyzed reports for ML (column (2)) rose in 2009 and 2010 and then dropped considerably. Also the ratio of dismissed reports for ML to reports received from intermediaries (column (3)) was higher in 2009 and

Table 4 Pattern of STRs, type-I and proxies of type-I error incidence (Italian provinces, 2009–2012); type-I, Italian provinces, 2008–2012

Year	[reports received from intermediaries ^(a) —analyzed reports for ML ^(b)]/reports received from intermediaries ^(a)	dismissed reports for ML ^(c) /analyzed reports for ML ^(b)	dismissed reports for ML ^(c) /reports received from intermediaries ^(a)
	(1)	(2)	(3)
2008	0.077	0.075	0.069
2009	0.124	0.207	0.181
2010	0.279	0.131	0.094
2011	0.374	0.040	0.025
2012	0.075	0.054	0.050

Data Source: Italian FIU. Tables from FIU Anti-money laundering notebooks Statistical data and annual reports

^(a)Reports received from financial intermediaries: all types of reports for suspicious transactions (money laundering, terrorism, arms trafficking)

^(b)Reports received from all reporting agents (financial intermediaries, professionals and non-financial operators, others) for money laundering

^(c)Dismissed reports for money laundering (financial intermediaries, professionals and non-financial operators, others)

2010 (compared to 2008) and then fell. Though, some considerations are worthwhile to explain its pattern.

In general, all proxies of type-I error, support the “crying wolf” hypothesis, at least in the period immediately after the introduction of the risk-based approach, showing a substantial increase in 2009–2010 (or 2011) compared to the pre-reform period in 2008. However, after reaching their maximum after two or three years, they suddenly dropped, stabilized, or even improved upon the pre-reform levels. Given the time required to process STRs, the initial backlog and the increase in the STRs’ dismissal rate, which we attribute to a decrease in the reporting test adopted by financial intermediaries seems to have occurred. This seems reasonable to the extent that the new regulation caused relevant changes in banks’ attitudes, which plausibly interpreted the cost of signaling as exempt from sanctions, or subject to relatively low reputational costs, along with the fact that the transposition of the new legislation led to difficulties in learning the new reporting procedures. This caused an anomalous wave of STR, which is also testified by the FIU Annual Report of 2010 (referring to year 2009), reporting that: “The length of reporting times often makes it difficult to promptly intercept suspicious flows, nullifying the preventive effectiveness of the system.” A similar judgment is given in the Report of the following year: “Despite the continuous increases in productivity achieved by the Unit in the financial analysis of STRs, the large and persistent growth in the number of incoming reports has determined an increase in the stock of STR waiting for processing” (FIU Annual Report of 2011, referring to year 2010).

Nevertheless, the following improvement and consequent (sudden) break occurring in our proxies of type-I error does not represent a failure of our model. The

enhancements that took place were instead motivated by an important technological change that occurred at the FIU. Indeed, the implementation of a new IT system for the acquisition and processing of reports has allowed the intelligence unit to cope with the exceptional increase, to accelerate and rationalize the screening process, as documented in the FIU Annual Report of 2012 (referring to 2011).

Finally, given that there are no reliable proxies for type-II errors, we exclusively rely on and draw inferences from the model (Fig. 5) to determine whether type-II errors increased during the period of interest. We cautiously infer that as the type-I error continuously increased until 2017, the opposite occurred with type-II errors. In conclusion, we feel reasonably confident in asserting that while the risk-based approach favored higher standards of deterrence in money laundering activities, pushing the reporting test toward the gold standard does not necessarily increase accuracy (see Fig. 5).

4 Concluding Remarks

The risk-based approach involves financial intermediaries in the AML system proactively. These actors must establish risk-based procedures and criteria (such as the KYC principle) to report certain transactions as suspicious to the FIU. From this perspective, the first level of the AML system can change the reporting test used to report a given transaction to the FIU.

The theoretical model proposed in this paper offers an interpretation of the behavior of the agents signaling STR to the FIU with respect to two important goals: the deterrence effects of their reporting activities and the accuracy of those reports. We empirically assessed the model's main predictions using multivariate techniques and sufficient statistics. The analysis, which was based on Italian data, focused on the role of financial intermediaries—the largest pool of actors submitting STRs to the FIU. We first tested the effects of the introduction of the risk-based methodology on deterrence and type-I errors by means of regression analysis (using disaggregated data at the provincial level) combined with aggregate statistics. Thereafter, we used the model's predictions to draw inferences regarding the pattern of type-II errors and, consequently, accuracy.

The empirical outcome suggests that, in the period of interest (i.e., in the years immediately following the introduction of the risk-based approach), financial intermediaries lowered the reporting test required to report a transaction as suspicious. The adoption of this “tougher” approach by the Italian intermediaries might have been motivated by the fact that these intermediaries are severely sanctioned if they do not report transactions subsequently detected as money-laundering transactions. The data combined with the model's predictions suggest that the reporting test moved closer to the gold standard, further promoting deterrence.

We found that the observed increase in STR activity in relation to total bank transactions could be explained by a decrease in the reporting test adopted by the intermediaries. The model predicts an increase in the incidence of type-I errors, which was confirmed by aggregate data on dismissal rates from the FIU. These findings might relate to the fact that intermediaries and professionals are not formally punished for

over-reporting. Conversely, we conclude that the incidence of type-II errors must have decreased during the period of interest—a conclusion that stems directly from the model. This is also motivated by an increase in reporting tests. In terms of policy, this inferred conclusion is particularly important, as data or proxies for the incidence of type-II errors are not easily observed at any layer of reporting activity.

Finally, although the risk-based approach aims to improve reporting quality, Italian authorities should be aware that although the MLR seems to be relatively stable (i.e., close to the maximum deterrence), the incidence of reporting errors is not at the minimum. In particular, while the inference about type-II errors is encouraging, type-I errors increased following the introduction of the new reporting rules, though largely compensated by substantial investment in IT-based screening technologies and the number of employees in the FIU.²⁴ However, these aspects could be detrimental to a bank’s reputation and client retention, while the AML authorities are doing most of the job of managing type-I errors.

From a policy perspective, the emerging evidence seems to indicate that the risk-based paradigm is helpful in combatting money laundering, at least in Italy. However, false positives remain a major issue that deserve further consideration. Researchers may wish to apply the approach proposed in this paper to other contexts to understand where different countries are positioned in terms of the accuracy of the information transmitted to the FIUs and the deterrence of money laundering activities.

Acknowledgements The authors are grateful to the editor and two anonymous referees, and to Mario Gara and staff at the Intelligence Financial Unit for their precious suggestions. They also thank the participants at the 2018 Annual Conference of the European Association of Law and Economics (EALE), the 2018 Annual Conference of the Italian Association of Law and Economics (SIDE-ISLE). They also kindly acknowledge the Baffi-Carefin Center at Bocconi University for financial support. All errors are the authors’ own. This study received funding from the PRIN Programme 2015 of the Italian Ministry of Education, University and Research (MIUR) for the Project “Bad Seeds: Diffusion and Effects of Criminal Organizations” [grant number IRIS: 2016-NAZ-0105].

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix 1

All the proofs in this Appendix are provided without recurring to the assumption of Normality of the two conditional probability distributions of the net evidence.

²⁴ Between 2009 and 2012, the number of employees increased from 97 to 121. Given the high professional level of the FIU’s employees, this increase is relevant (FIU, Annual Reports, 2010–2012).

The Gold Standard

For a given transaction, the most accurate reporting test \hat{e} is such that the true positive rate $1 - G_E(\hat{e})$ is maximized while the false positive rate $1 - I_E(\hat{e})$ is minimized.

$$\begin{cases} \max_{\hat{e}} 1 - G_E(\hat{e}) \\ \min_{\hat{e}} 1 - I_E(\hat{e}) \end{cases}$$

Resolving for the first-order conditions and checking for the second-order conditions, we conclude that, for a given transaction, the most accurate test $\hat{e}_{gold} : i_E(\hat{e}_{gold}) = g_E(\hat{e}_{gold})$, where the conditional probability distributions cross.

It is worth noting that the equilibrium condition $i_E(\hat{e}_{gold}) = g_E(\hat{e}_{gold})$ does not require any assumption of symmetry of the two distributions.

The Money Laundering Rate Negatively Depends on Sanctions

$$MLR(S, \hat{e}) = 1 - S[I_E(\hat{e}) - G_E(\hat{e})]$$

$\frac{\partial MLR(S, \hat{e})}{\partial S} = -[I_E(\hat{e}) - G_E(\hat{e})] < 0$ (because of the first-order stochastic dominance).

The Money Laundering Rate is Minimized by the Gold Standard \hat{e}_{gold}

Let define $\hat{e}_{gold} : i_E(\hat{e}_{gold}) = g_E(\hat{e}_{gold})$ and study $\frac{\partial MLR(S, \hat{e})}{\partial \hat{e}} = -S[i_E(\hat{e}) - g_E(\hat{e})]$. We observe the following:

$$\begin{aligned} \frac{\partial MLR(S, \hat{e})}{\partial \hat{e}} &< 0, \quad \text{for } -\infty < \hat{e} < \hat{e}_{gold} \\ \frac{\partial MLR(S, \hat{e})}{\partial \hat{e}} &= 0, \quad \text{for } \hat{e} = \hat{e}_{gold} \\ \frac{\partial MLR(S, \hat{e})}{\partial \hat{e}} &> 0, \quad \text{for } \hat{e}_{gold} < \hat{e} < \infty \end{aligned}$$

Study of the Reporting Activity $STR(S, \hat{e})$

$$STR(S, \hat{e}) = (1 - S[I_E(\hat{e}) - G_E(\hat{e})])(1 - G_E(\hat{e})) + S[I_E(\hat{e}) - G_E(\hat{e})](1 - I_E(\hat{e}))$$

$$\frac{\partial STR(S, \hat{e})}{\partial S} = -[I_E(\hat{e}) - G_E(\hat{e})]^2 \leq 0 (= 0 \text{ only for } I_E(\hat{e}) = G_E(\hat{e}))$$

Recalling that sanctions S are nonnegative, the standard behaviour of probability functions $i_E(\hat{e})$ and $g_E(\hat{e})$ and density functions $I_E(\hat{e})$ and $G_E(\hat{e})$ and that

$i_E(\hat{e}) - g_E(\hat{e}) > 0$ for $-\infty < \hat{e} < \hat{e}_{gold}$
 $(I_E(\hat{e}) - G_E(\hat{e})) \geq 0$ and $i_E(\hat{e}) - g_E(\hat{e}) = 0$, for $\hat{e} = \hat{e}_{gold}$ by assumption,
 $i_E(\hat{e}) - g_E(\hat{e}) < 0$, for $\hat{e}_{gold} < \hat{e} < \infty$
 tions, we can observe the following:

$$\begin{aligned} STR(S, \hat{e}) &= 1 \text{ for } \hat{e} : I_E(\hat{e}) = G_E(\hat{e}) \rightarrow 0 \\ STR(S, \hat{e}) &= 0 \text{ for } \hat{e} : I_E(\hat{e}) = G_E(\hat{e}) \rightarrow 1 \end{aligned}$$

Moreover, since $\frac{\partial STR(S, \hat{e})}{\partial \hat{e}} = - \left[\underbrace{2S}_{\geq 0} \underbrace{(i_E(\hat{e}) - g_E(\hat{e}))}_{\geq} \underbrace{(I_E(\hat{e}) - G_E(\hat{e}))}_{\geq 0} + \underbrace{g_E(\hat{e})}_{\geq 0} \right],$

$$\frac{\partial STR(S, \hat{e})}{\partial \hat{e}} < 0 \text{ for } -\infty < \hat{e} \leq \hat{e}_{gold}$$

$$\frac{\partial STR(S, \hat{e})}{\partial \hat{e}} \begin{matrix} \leq \\ > \end{matrix} \text{ for } \hat{e} > \hat{e}_{gold}$$

$$\frac{\partial STR(S, \hat{e})}{\partial \hat{e}} \leq 0 \text{ for } \hat{e} \gg \hat{e}_{gold}$$

$$\frac{\partial STR(S, \hat{e})}{\partial S} = -[I_E(\hat{e}) - G_E(\hat{e})]^2 \leq 0 (= 0 \text{ only for } I_E(\hat{e}) = G_E(\hat{e}))$$

Study of the Reporting Errors $ER(S, \hat{e})$

$$ER(\hat{e}) = S[I_E(\hat{e}) - G_E(\hat{e})](1 - I_E(\hat{e})) + (1 - S[I_E(\hat{e}) - G_E(\hat{e})])G_E(\hat{e})$$

$$\frac{\partial ER(S, \hat{e})}{\partial S} = [I_E(\hat{e}) - G_E(\hat{e})](1 - I_E(\hat{e}) - G_E(\hat{e}))$$

$$\frac{\partial ER(S, \hat{e})}{\partial S} > 0, \text{ for } \hat{e} < \hat{e}_{gold}$$

$$\frac{\partial ER(S, \hat{e})}{\partial S} = 0, \text{ for } \hat{e} = \hat{e}_{gold}$$

$$\frac{\partial ER(S, \hat{e})}{\partial S} < 0, \text{ for } \hat{e} > \hat{e}_{gold}$$

$$\begin{aligned} \frac{\partial ER(S, \hat{e})}{\partial \hat{e}} &= S[i_E(\hat{e}) - g_E(\hat{e})](1 - I_E(\hat{e}) - G_E(\hat{e})) \\ &\quad - S[I_E(\hat{e}) - G_E(\hat{e})](i_E(\hat{e}) + g_E(\hat{e})) + g_E(\hat{e}) \end{aligned}$$

$$\begin{aligned} \frac{\partial ER(S, \hat{e})^2}{\partial^2 \hat{e}} \Big|_{\hat{e}_{gold}} &= S[-(i_E(\hat{e}) - g_E(\hat{e}))(i_E(\hat{e}) + g_E(\hat{e})) \\ &\quad + \left(\frac{\partial i_E(\hat{e})}{\partial \hat{e}} - \frac{\partial g_E(\hat{e})}{\partial \hat{e}} \right) (1 - I_E(\hat{e}) - G_E(\hat{e}))] \\ &\quad - S[(i_E(\hat{e}) - g_E(\hat{e}))(i_E(\hat{e}) + g_E(\hat{e})) \\ &\quad + [I_E(\hat{e}) - G_E(\hat{e})] \left(\frac{\partial i_E(\hat{e})}{\partial \hat{e}} + \frac{\partial g_E(\hat{e})}{\partial \hat{e}} \right)] + \frac{\partial g_E(\hat{e})}{\partial \hat{e}} \end{aligned}$$

Note that the gold standard \hat{e}_{gold} corresponds to a stationary state only if $\frac{\partial ER(S, \hat{e})}{\partial \hat{e}} \Big|_{\hat{e}_{gold}} = 0$. Recalling that $i_E(\hat{e}_{gold}) = g_E(\hat{e}_{gold})$, we easily verify that $\frac{\partial ER(S, \hat{e})}{\partial \hat{e}} \Big|_{\hat{e}_{gold}} = -S[I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold})]2g_E(\hat{e}_{gold}) + g_E(\hat{e}_{gold}) = 0$ iff $S = \frac{g_E(\hat{e}_{gold})}{(I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold}))}$.

In addition, recalling that $\frac{\partial i_E(\hat{e})}{\partial \hat{e}}|_{\hat{e}_{gold}} = -\frac{\partial g_E(\hat{e})}{\partial \hat{e}}|_{\hat{e}_{gold}}$, we conclude that $\frac{\partial ER(S, \hat{e})^2}{\partial \hat{e}}|_{\hat{e}_{gold}} = \frac{\partial g_E(\hat{e}_{gold})}{\partial \hat{e}_{gold}} > 0$. Therefore, when in \hat{e}_{gold} there is a stationary state, it is a minimum. This implies that, when $S = \frac{0.5}{(I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold}))}$, $ER(S, \hat{e})$ is locally minimized by setting $\hat{e} = \hat{e}_{gold}$.

However, when $S \neq \frac{0.5}{(I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold}))}$, \hat{e}_{gold} does not minimize the error rate. In particular, when $S < \frac{0.5}{(I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold}))}$, $\frac{\partial ER(S, \hat{e})}{\partial \hat{e}}|_{\hat{e}_{gold}} > 0$; this means that the local minimum of the error rate²⁵ is on the left of \hat{e}_{gold} ; the smaller the sanctions more on the left the local minimum of the errors rate. Vice versa, when $S > \frac{0.5}{(I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold}))}$, $\frac{\partial ER(S, \hat{e})}{\partial \hat{e}}|_{\hat{e}_{gold}} < 0$; this means that the local minimum of the error rate is on the right of \hat{e}_{gold} ; the higher the sanctions more on the right the local minimum of the errors rate.

Defining \hat{e}_{minER} as the reporting test that minimizes the error rate $\rightarrow \hat{e}_{minER} = \hat{e} : \min_{\hat{e}} ER(S, \hat{e})$ and summarizing the reasoning above, we conclude that:

$$\begin{aligned} \text{if } S &< \frac{0.5}{I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold})}, \hat{e}_{minER} < \hat{e}_{gold} \\ \text{if } S &= \frac{0.5}{I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold})}, \hat{e}_{minER} = \hat{e}_{gold} \\ \text{if } S &> \frac{0.5}{I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold})}, \hat{e}_{minER} > \hat{e}_{gold} \end{aligned} \quad (\text{A.1})$$

Finally, recalling that $MLR(S, \hat{e}) = 1 - S[I_E(\hat{e}) - G_E(\hat{e})]$, we observe that $MLR(S, \hat{e}_{gold}) = 1 - S[I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold})] \equiv MLR_{min}$ and then that $S = \frac{1 - MLR_{min}}{I_E(\hat{e}_{gold}) - G_E(\hat{e}_{gold})}$. Therefore, given the relation between sanctions and the minimum money laundering rate, we can also rewrite (A.1) as:

$$\begin{aligned} \text{if } MLR_{min} &> 0.5, \hat{e}_{minER} < \hat{e}_{gold} \\ \text{if } MLR_{min} &= 0.5, \hat{e}_{minER} = \hat{e}_{gold} \\ \text{if } MLR_{min} &< 0.5, \hat{e}_{minER} > \hat{e}_{gold} \end{aligned} \quad (\text{A.2})$$

Moreover, by following the same reasoning made for sanctions, we conclude that the higher the $MLR_{min} > 0.5$, more on the left the local minimum of the errors rate; the lower the $MLR_{min} < 0.5$, more on the right the local minimum of the errors rate.

Appendix 2

See Table 5

²⁵ Technically, for some parameters, it might be an inflection point and not a local minimum.

Table 5 Summary statistics, Italian provinces, 2009–2012

Variable	Obs	Mean	Std. Dev	Min	Max
STR (number per 100.000 inhabitants) (¹)	412	59.69	48.38	0	391
ML police reports (number per 100.000 inhabitants) (²)	412	2.09	2.14	0	15
Per capita GDP (real value added at 2010 price level—mil.EUR)	412	0.02	0.01	0.01	0.05
Duration of criminal trials (days)	412	327.41	95.73	133	588
Detected ML crimes/pop (ML crimes with known offender /total ML crimes)	412	20.52	15.4	0	114
Number of bank branches per 100.000 inhabitants	412	58.88	19.75	21	105
Cash inflows to the banking system per 100.000 inhabitants	412	143.03	40.86	50	443
Number of real estate transactions per 100.000 inhabitants	412	1673	404	763	3053
Number of real estate loans per 100.000 inhabitants	412	647	278	83	1525
Predicate crimes (per 100.000 inhabitants except (³):					
Microcrime index (³)	412	12.35	4.59	3	280
Exploitation and abetting of prostitution	412	2.56	2.08	0	18
Racketeering	412	9.49	4.87	0	31
Frauds	412	154.97	45.78	30	356
Counterfeiting	412	15.5	14.72	0	98
Drug trafficking	412	50.42	21.35	0	144
Armed robbery	412	36.59	35.22	0	284
Criminal association	412	1.34	1.85	0	18
Mafia-type association	412	0.16	0.65	0	9
Usury	412	0.62	0.66	0	4
Corruption	412	4.43	3.16	0	22
Tax_gap (at municipality level) (³)	412	69.38	7.10	45	83
Tax_gap (at provincial level) (³)	412	71.47	11.17	15	96

Notes. Obs. 412

(1) Total inflows in the 12 months prior to the observation

(2) Police reports refer to the number of crimes reported to the Police Authority

(3) Percentage

Appendix 3

See Table 6.

Table 6 Factor loadings and unique variances, Italian provinces, 2009–2012

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Uniqueness
Microcrime	0.85	-0.16	0.12	-0.06	0.03	-0.01	0.12	0.21
Prostitution	0.31	0.27	0.31	0.11	0.03	0.01	-0.02	0.72
Racketeering	0.62	0.42	-0.09	-0.13	-0.11	0.02	-0.04	0.40
Frauds	0.94	-0.08	-0.03	-0.07	0.00	0.08	-0.06	0.10
Counterfeiting	0.51	0.43	-0.05	0.21	-0.05	-0.07	0.04	0.50
Drug trafficking	0.87	0.16	0.12	0.12	0.00	0.01	-0.06	0.19
Armed robbery	0.57	0.43	0.06	-0.22	-0.06	0.00	0.05	0.43
Criminal association	0.37	0.12	0.07	-0.03	0.21	-0.07	-0.03	0.80
Mafia-type association	0.07	-0.37	0.17	0.10	-0.14	0.09	0.02	0.79
Usury	0.26	-0.05	-0.07	0.01	0.25	0.10	0.02	0.85
Corruption	0.41	0.34	-0.26	0.15	0.00	0.06	0.04	0.62
Tax_gap (at municipality level)	0.82	-0.54	-0.06	0.00	-0.01	-0.06	-0.02	0.02
Tax_gap (at provincial level)	0.83	-0.53	-0.10	0.02	-0.03	-0.04	0.00	0.03

References

- Abadinsky H (2010) Organized crime, 9th edn. Wadsworth, Cengage Learning, Belmont
- Arnone M, Borlini L (2010) International anti-money laundering programs. *J Money Launde Control* 13(3):226–271. <https://doi.org/10.1108/13685201011057136>
- Axelrod RM (2017) Criminality and suspicious activity reports. *J Fin Crime* 24(3):461–471. <https://doi.org/10.1108/JFC-03-2017-0019>
- Barone R, Masciandaro D (2019) Cryptocurrency or usury? Crime and alternative money laundering techniques. *Eur J Law Econ* 47(2):233–254. <https://doi.org/10.1007/s10657-019-09609-6>
- Black J, Baldwin R (2010) Really responsive risk-based regulation. *Law Policy* 32(2):181–213. <https://doi.org/10.1111/j.1467-9930.2010.00318.x>
- Chetty R (2009) Sufficient statistics for welfare analysis: a bridge between structural and reduced-form methods. *Annu Rev Econ* 1(1):451–488. <https://doi.org/10.1146/annurev.economics.050708.142910>
- Dalla Pellegrina L, Masciandaro D (2009) The risk-based approach in the new european anti-money laundering legislation: a law and economics view. *Rev Law Econ* 5(2):931–952
- Dalla Pellegrina L, Di Maio G, Masciandaro D, Saraceno M (2020a) Organized crime, suspicious transaction reporting and anti-money laundering regulation. *Reg Stud* 54(12):1761–1775. <https://doi.org/10.1080/00343404.2020.1772963>
- Dalla Pellegrina L, Di Maio G, Saraceno M (2020b) Standard of Proof, Accuracy and Deterrence. Mimeo.
- Demetis DS (2010) Technology and anti-money laundering: a systems theory and risk-based approach. Edward Elgar, Cheltenham
- Draghi M (2007) The prevention and suppression of money laundering—Testimony at the Parliamentary Committee of Inquiry into Mafia-like Criminal Organizations. Retrieved from <http://www.bis.org/review/r070727c.pdf>
- Ferwerda J, van Saase A, Unger B, Getzner M (2020) Estimating money laundering flows with a gravity model-based simulation. *Sci Rep* 10(1):1–11
- Gara M, Manaresi F, Marchetti DJ, Marinucci M (2019) The impact of anti-money laundering oversight on banks’ suspicious transaction reporting: evidence from Italy. *Bank Italy Occas Pap* 491.
- Gara M, Pauselli C (2015) Looking at ‘Crying wolf’ from a different perspective: an attempt at detecting banks under- and over-reporting suspicious transactions. *Banca D’italia - Quaderni Dell’ anticiclaggio* 4:1–26
- Imanpour M, Rosenkranz S, Westbrook B, Unger B, Ferwerda J (2019) A microeconomic foundation for optimal money laundering policies. *Int Rev Law Econ* 60:105856. <https://doi.org/10.1016/j.irl.2019.105856>
- Jayasekara SD (2020) How effective are the current global standards in combating money laundering and terrorist financing? *J Money Laund Control*.
- Jeans N (2016) Risk-Based Approach to KYC: Sound concept, complex reality Statement of intent. Retrieved from <https://risk.thomsonreuters.com/en/resources/white-paper/risk-based-approach-kyc.html>
- Jöreskog KG (1969) A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34(2):183–202. <https://doi.org/10.1007/BF02289343>
- Lowe RJ (2017) Anti-money laundering—the need for intelligence. *J Fin Crime* 24(3):472–479. <https://doi.org/10.1108/JFC-04-2017-0030>
- Marzban C (2004) The ROC curve and the area under it as performance measures. *Weather Forecast* 19:1106–1114
- Mugarura N (2011) The institutional framework against money laundering and its underlying predicate crimes. *J Fin Regul Comp*.
- Rizzolli M, Saraceno M (2013) Better that ten guilty persons escape: punishment costs explain the standard of evidence. *Public Choice* 155(3–4):395–411. <https://doi.org/10.1007/s11127-011-9867-y>
- Takáts E (2011) A theory of “crying wolf”: the economics of money laundering enforcement. *J Law Econ Organ* 27(1):32–78. <https://doi.org/10.1093/jleo/ewp018>
- Unger B, van Waarden F (2009) How to dodge drowning in data? Rule- and risk-based anti money laundering policies compared. *Rev Law Econ* 5(2):953–985. <https://doi.org/10.2202/1555-5879.1423>
- Zou KH, O’Malley AJ, Mauri L (2007) Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* 115:654–657