BANCA D'ITALIA
EUROSISTEMA

UIF
Unità di Informazione Finanziaria per l'Italia

Quaderni dell'antiriciclaggio

Analisi e studi

The size of money laundering and other illicit financial conduct for Italy

Michele Giammatteo

gennaio 2025

numero 26

BANCA D'ITALIA
EUROSISTEMA

UIF
Unità di Informazione Finanziaria per l'Italia

Quaderni dell'antiriciclaggio

Analisi e studi

The size of money laundering and other illicit financial conduct for Italy

Michele Giammatteo

n. 26 – gennaio 2025

La collana Quaderni dell'antiriciclaggio *ha la finalità di presentare statistiche, studi e documentazione su aspetti rilevanti per i compiti istituzionali dell'Unità d'Informazione Finanziaria per l'Italia.*

La collana si articola in diversi filoni: il filone Statistiche *presenta, con periodicità semestrale, statistiche sulle segnalazioni ricevute e dati sulle attività dell'Unità; il filone* Rassegna normativa *illustra i principali aggiornamenti della normativa e della giurisprudenza in materia AML/CFT; il filone* Analisi e studi *comprende contributi sulle tematiche e sui metodi in materia di contrasto al riciclaggio e al finanziamento del terrorismo. I lavori pubblicati riflettono esclusivamente le opinioni degli autori, senza impegnare la responsabilità delle Istituzioni di appartenenza.*

Unità di Informazione Finanziaria per l'Italia
Per la pubblicazione cartacea: autorizzazione del Tribunale di Roma n. 1942013 del 30 luglio 2013
Per la pubblicazione telematica: autorizzazione del Tribunale di Roma n. 1932013 del 30 luglio 2013

Direttore responsabile
Enzo Serata

Indirizzo
Largo Bastia, 35 – 00181 Roma – Italia

Telefono
+39 0647921

Sito internet
https://uif.bancaditalia.it/

ISSN 2283-3498 (stampa)
ISSN 2283-6977 (online)

Stampato nel mese di gennaio 2025

Grafica e stampa a cura della Divisione Editoria e stampa della Banca d'Italia

THE SIZE OF MONEY LAUNDERING AND OTHER ILLICIT FINANCIAL CONDUCT FOR ITALY

by Michele Giammatteo*

Abstract

Estimating the size of money laundering (ML) is inherently challenging due to its hidden nature and the diverse techniques employed by launderers. The effectiveness of measures to fight ML varies across countries, and the precise scale of it remains elusive. Existing literature provides empirical analyses that are often hampered by data and methodological shortcomings, low accuracy, and a lack of replicability. This study leverages a unique dataset from the Italian Financial Intelligence Unit (UIF) to estimate ML in Italy from 2018 to 2022, employing a novel three-phase empirical approach: 1) selecting relevant Suspicious Transaction Reports (STRs), 2) applying a machine learning algorithm to identify reliable transaction values, and 3) imputing unreliable data with more reliable observations. The results provide a conservative estimate of ML, averaging 1.8% of Italy's GDP during the period of analysis. Accurately estimating ML is crucial for raising awareness, identifying trends, and targeting AML policies.

Sommario

Stimare il valore del riciclaggio è un compito complesso a causa della sua natura occulta e delle svariate tecniche utilizzate per mascherare l'origine illecita dei proventi. L'efficacia delle misure di contrasto varia da paese a paese e il valore del riciclaggio è spesso un dato sconosciuto. La letteratura esistente offre analisi empiriche che spesso sono limitate da carenze di dati e metodologiche, bassa accuratezza e mancanza di replicabilità. Partendo da dati aggregati ricavati dalle Segnalazioni di Operazioni Sospette (SOS) ricevute dalla UIF, questo studio propone una stima del valore dei flussi finanziari coinvolti nel riciclaggio e nelle altre condotte finanziarie illecite in Italia nel periodo 2018-2022. La metodologia adottata si articola nelle seguenti tre fasi: 1) selezione accurata delle SOS rilevanti sotto il profilo finanziario o investigativo; 2) applicazione di un algoritmo di *machine learning* per identificare i dati "affidabili"; 3) sostituzione dei dati "non affidabili" con quelli delle unità statisticamente più simili tra le osservazioni ritenute "affidabili". I risultati forniscono una stima conservativa del valore del riciclaggio e delle altre condotte illecite in Italia che, per il periodo analizzato, è pari in media all'1,8% del PIL. Una valutazione affidabile della dimensione dei flussi finanziari oggetto di riciclaggio rappresenta uno strumento fondamentale per aumentare la consapevolezza di questo fenomeno, monitorarne le tendenze, oltre che per orientare le politiche di contrasto.

JEL Classification: C49, G28, K42.
Keywords: Money laundering, Suspicious Transaction Reports, Italy, Quantile Random Forest, Imputation.

---

* Bank of Italy, Financial Intelligence Unit, Information Exploitation and Technological Innovation Directorate, Financial Flows Analysis Division.

Contents

1. Introduction[1]

Estimating the scale of money laundering (ML) is an ambitious and challenging endeavor. First, ML is a complex illicit activity that, by its very nature, is deliberately concealed. Second, professional money launderers deploy a wide range of sophisticated techniques specifically designed to disguise the illegal origin of the funds they transfer both within and across borders, making these origins difficult to trace. This poses a series of challenges for institutions tasked with preventing and prosecuting ML. To address these challenges, these institutions implement comprehensive, innovative and technologically advanced financial analysis and investigative techniques.

As it is well known, the effectiveness of these efforts varies significantly across countries. Some countries face higher risk due to weak anti-money laundering (AML) regulations, political instability, or proximity to regions with heightened exposure to illicit financial flows. On the other hand, many countries have successfully implemented robust laws and institutional frameworks to counter ML more effectively. However, a common challenge persists across both groups: the true scale of ML remains largely unknown.

Research available in the literature – whether published by academia, conducted by international organizations, or documented in government reports – provides a moderate range of empirical analyses regarding the scale of ML and the theoretical frameworks on which they are based. Moreover, these findings often yield limited and narrow results. Sometimes the estimates rely on broad global approximations or conjectures, which are provided sporadically rather than on a regular basis. They occasionally refer not to ML per se, but to related phenomena, such as the extent of illegal proceeds (often associated with a single crime, such as drug trafficking) or other components of the non-observed economy. At times, estimates are derived using *ad hoc* data and information on criminal proceeds estimation (UNODC, 2011), expert opinions (Camdessus, 1998) or undocumented exercises of data extrapolation over time and among countries. According to these sources, the global amount of money laundered each year is estimated to be approximately 2-5% of global GDP.

In a few studies, robust empirical methodologies – mostly belonging to the class of indirect methods – are applied. Among them, it is worth mentioning the Currency Demand approach (Tanzi, 1996; Ardizzi et al., 2014), the Gravity Models (Walker, 1995, 1999, 2007; Walker and Unger, 2009), the Multiple Indicators – Multiple Causes method (Schneider, 2007; Schneider, 2010), and the development and numerical simulation of theoretical dynamic general equilibrium models (Bagella et al., 2009; Barone and Masciandaro, 2011). These methodologies are applied at times at the country level or with reference to supranational territorial entities such as the EU, OECD countries, or the entire world.

Surprisingly, nearly all the available empirical research neglects the information on Suspicious Transaction Reports (STRs), which are sent to Financial Intelligence Units by obligated entities to comply with AML obligations.[2] The only two exceptions (to our knowledge) can be found in the works of Walker (1995) and Ferwerda et al. (2020). The former contribution primarily aims to justify their limited use by acknowledging that "[STRs] *do not measure money laundering directly or completely. (…) some apparently suspect transactions turn out, on investigation, to be legitimate, and some suspect transactions are a very thin end of a very large wedge*." (p. 33). On the contrary, Ferwerda et al. (2020) use a dataset of financial transactions suspected of ML, made available by the FIU of the Netherlands. They employ a direct econometric gravity model

[1] The views expressed in this paper are those of the author and do not involve the responsibility of UIF or the Bank of Italy. The usual disclaimers apply. I am grateful to Mario Gara, Stefano Iezzi, Andrea Silvestrini, and seminar participants at UIF. Special thanks to Francesco Fiorini and Francesco Tondi (UIF) for their valuable contributions to the key phase of selection and interpretation of STRs data; Luciano Cavalli (ISTAT) for providing data on the shadow added value at province level on behalf of the working group on "Predictive Use of Payment System Transaction Data" of the ISTAT-Bank of Italy Coordination Committee for cooperation in research and the exchange of statistical information (Ardizzi and Righi, 2022).

[2] A major motivation may stem from restrictions on the dissemination of AML data due to confidentiality requirements, although this obstacle could potentially be overcome with the use of sufficiently aggregated and anonymized data.

estimation and provide figures on the magnitude of ML for many countries worldwide.[3] Although their methodology differs from that employed here, their approach starts from a fundamental premise that we fully agree with: the most accurate data on ML are those collected for AML purposes. While such data may be affected by under- and over-reporting, they still represent the most reliable and relevant measure of the phenomenon. By carefully selecting these data and employing appropriate statistical methods, it is possible to reasonably offset a significant portion of the bias that may affect the estimation process.

The contribution of this work to the existing literature is twofold. Firstly, a unique database containing data on ML activity is used, to the best of our knowledge, for the first time. This database combines information from STRs[4] with other valuable financial data available at the Italian Financial Intelligence Unit (UIF), as well as confidential data provided by the Italian National Institute of Statistics (ISTAT). It also includes publicly available data from the websites of ISTAT and the Italian Revenue Agency ("*Agenzia delle Entrate*"). Secondly, we develop an original empirical approach for estimating the size of ML in Italy, consisting of the following three main phases:

1. **Data Selection and Database Creation**: First, we employ a rigorous selection of relevant STRs to build a database of "observed ML" that closely aligns with the phenomenon under analysis. In this regard, we only consider STRs that, based on financial analyses conducted by UIF experts, are either classified as high-risk or have received significant feedback from enforcement agencies.
2. **Estimation Using Machine Learning**: Secondly, we use a machine learning algorithm to estimate the value of ML financial transactions that each bank in each of the Italian provinces is expected to report to the UIF on the basis of a set of explanatory variables related to ML. By comparing these estimates with observed STRs, we identify two sets of bank-province observations: those featuring a value of ML roughly in line with the expected value, thus considered as "reliable" observations, and those that are inconsistent with respect to it (either less or in excess), and may thus be regarded as "unreliable" observations.
3. **Imputation with Predictive Mean Matching**: Finally, we apply a version of the Predictive Mean Matching method for imputation, as developed by Little (1988) on the basis of the fundamental contribution to statistical matching of Rubin (1986). This approach is used to replace the values of ML financial transactions reported by "unreliable" bank-province observations (i.e., the output of phase 2) with those from the most statistically similar "reliable" bank-provinces, based on an appropriate distance function.

This work estimates the size of ML in Italy over a five-year period (2018-2022), with baseline results indicating that ML accounted for an average of 1.8% of GDP. This percentage is broadly consistent with ISTAT's official estimates of illegal activities, which accounted for 1.0% of Italian GDP[5] in 2021. The difference can be explained as follows: ISTAT's estimates cover only three types of illegal activities – drug trade, prostitution, and cigarette smuggling – that, according to Transcrime (2015), represent only half of the total illicit revenues. As a result, all illegal activities together should account for approximately 2% of Italy's GDP (Mocetti and Rizzica, 2023). Therefore, the ML estimate we obtained is encouragingly lower than the estimated value of predicate crimes it encompasses.[6]

Estimating the size of ML is of paramount importance as it raises awareness about the phenomenon, identifies trends and patterns, breaks down its amount at geographical levels, and provides risk indicators useful for establishing a proper risk-based approach in AML action. It also potentially

[3] Some countries are excluded if they have missing data for some of the variables used in the estimation procedure.

[4] The Legislative Decree 231/2007 requires financial intermediaries, other persons engaging in financial activity, and a series of persons engaged in other (non-financial) activities to send to the UIF a suspicious transaction report "whenever they know, suspect or have reason to suspect that money-laundering or terrorist financing is being or has been carried out or attempted."

[5] See the last ISTAT report on Italian Non-Observed Economy (available only in Italian): https://www.istat.it/wp-content/uploads/2023/10/Report-ECONOMIA-NON-OSSERVATA-2021.pdf.

[6] A further important source of illegal proceeds relevant for ML comes from tax evasion. According to EU national accounting standards (ESA 2010), ISTAT includes it in the shadow economy component (namely, the underreporting of value-added) that, together with illegal activities, constitutes the overall aggregate of the Non-Observed Economy.

enables the evaluation of the impact of regulatory changes and other institutional developments in AML efforts on the scale of ML.

The rest of the paper is structured as follows. Section 2 discusses the feasibility and usefulness of estimating the magnitude of ML: it addresses the challenges and benefits of such estimations, considering factors like data availability, methodological robustness, and the implications of having accurate estimates for policy-making and law enforcement. Section 3 reviews the existing literature on empirical methods, highlighting various approaches, their strengths and weaknesses, and providing useful background for the current study. Section 4 outlines the empirical approach adopted in this paper for measuring ML in Italy: it describes in detail the methodological phases implemented and the data sources used in the estimation process. Section 5 presents the main findings of the data analysis, including estimates of the size of ML in Italy broken down by time periods and Italian regions. Finally, Section 6 summarizes the main findings of the paper, also addressing limitations of the study and proposing possible further developments.

2. Are the estimates of the volume of money laundering either feasible or useful?

The title of this section is borrowed from Peter Reuter's chapter in the Research Handbook of Money Laundering (Reuter, 2013). The final negative conclusion he reached is summarised in the following statement: "*My view is that knowing how much money is laundered serves no important policy purpose. It is simply one of those ornaments for conversations about the phenomenon. That is very fortunate since we have no methodology that plausibly would produce credible numbers.*" (p. 224). Two overarching thoughts emerge from this sentence: first, the quantification of ML is seen as non-essential, since AML systems generally operate independently of precise measurements; second, this primarily stems, among other factors, from the lack of reliable and persuasive methodologies for accurate estimation. Reuter (2013) further contends that there is little empirical evidence to suggest that ML directly destabilizes financial systems. In the rare instances where adverse consequences have been documented,[7] it is the underlying criminal activities associated with ML that cause harm, albeit predominantly impacting the individuals involved (admittedly also heavily) rather than the financial system at large.

More importantly, Reuter (2013) emphasizes that the information about the size of national ML is irrelevant for the agencies directly engaged in prevention and contrast activities. Indeed, none of these authorities consider the reduction of ML amount as a measure of outcome performance, nor is it explicitly required by any international standard. The main rationale behind this view lies in the substantial variability of ML, which arises from the diverse nature of illicit activities, varying in scope, the intensity of socio-economic consequences, and extent of individuals potentially involved. Setting specific targets to reduce ML volume could lead to unintended consequences, such as prioritizing less serious crimes with higher ML intensity over more severe threats, like terrorisms, simply to improve performance metrics. This could also potentially result in inconsistent practices across countries and distortions in the global AML system. Last but not the least, his key conclusion is reinforced by the paucity of available empirical evidence, either strongly subject to weak data and hypotheses or not implementing a sufficiently rigorous methodological approach.[8]

While one might agree with the logic of Reuter's reasoning, several objections could be raised. In particular, regarding the utility of estimating ML size, it is worth to acknowledge that as for every socio-economic phenomenon the process of knowledge achievement on its nature, trends, changes and interaction with other variables can be hardly pursued without the availability of accurate measurement or – by recurring to a probabilistic concept – statistical estimate. Of course, this can only follow the

[7] He explicitly cites the case of Latvia in the 1990s and of the Dominican Republic in 2002.
[8] "*What I hope to persuade you of is that there is no prospect, either in surveys of experts or in studies of crimes themselves as reflected in criminal justice statistics, for developing persuasive estimates.*" (Reuter, 2013; p. 224).

development of an underlying theory which conceptualizes the proper framework and relationships with other relevant phenomena.

In our view, the quantification of the size of ML represents an essential step for gauging its direct and indirect impacts on macroeconomic aggregates, as well as evaluating the effectiveness of prevention and contrast actions. The literature supporting this thought is abundant and diverse. Among the negative economic consequences of ML,[9] attention is drawn to potential financial instability resulting from unexpected variations in money demand unrelated to economic fundamentals, risk of integrity of the financial system that criminal organizations may seek to directly influence, heightened volatility of international capital flows and exchange rates, as well as general inefficiency of financial markets stemming from criminals' investment strategies which typically diverge from rational income maximization (Walker and Unger, 2009; Reuter, 2013). De Carolis et al. (2014) find that higher criminal density raises the cost of credit for businesses, induces greater demand for guarantees from banks, and implies potential negative effects on investment and economic growth. Novaro et al. (2022) provide evidence that entrepreneurial crimes (such as drug dealing, receiving stolen goods, and prostitution) positively affect housing market prices through increased demand. This finding is crucial not only for a matter of efficiency but also for equity, since differently from other 'richer markets' used for laundering money (like that of diamonds) inflated rises in residential housing are much more important in terms of social welfare losses and widening inequalities, being housing a typical destination of savings for all households.

While we fully acknowledge the economic relevance of understanding the size of ML, we also believe it should be of primary importance for scholars and practitioners in relevant research areas. Indeed, within every AML framework, the state-of-the-art entails adopting a risk-based approach to prevent and combat ML. Given limited resources, prioritizing and targeting efforts towards riskier sectors could facilitate the setting of priorities and provide guidance to agencies and institutions tasked with AML controls, both on an operational and strategic ground. In this context, estimating ML, at the national level or preferably at sub-national or bank level, would serve as a direct indicator of the financial system's exposure to ML risks.

In conclusion, although any estimate comes with some margin of error, it still provides a necessary indication of the considerable importance of ML (Schneider, 2013). At the same time, "*making statements about the seriousness of a problem and demonstrating its real size are two different issues*" (van Duyne, 1994, p. 59). Accurate estimates cannot be derived without reliable data and the adoption of rigorous and transparent methodologies.

3. Literature review

Estimating the size of ML is a primary research objective with important economic, social, and policy implications. It allows for comparing its potential socio-economic impact with that of other illegal phenomena and serves as crucial information for optimal allocation of public resources towards effective and efficient prevention and contrast activity. Therefore, it is surprising that data analysis devoted to ML size estimation has played only a marginal role in the assessment of national AML systems.

One possible explanation for this might rely on the numerous problems that arise with its definition, conceptual references, available data, and robust methodologies to support the estimation procedure. Regarding the definition, much effort has been made by supranational organizations (such as FATF and the Egmont Group) to harmonize it, but substantial differences still persist today since predicate offences, interpretation of jurisprudential orientations, and principles vary among

[9] The sole positive effects of ML may be identified in a possible increase in demand for goods and services, which in turn might imply positive growth of the legal economy. However, these short-term benefits would be more than compensated by the negative indirect consequences connected to the weakening of social, political and economic institutions.

jurisdictions.[10] To overcome this issue, there are two possible approaches: focussing on in-depth analysis of financial operations strictly aimed at concealing and integrating criminal proceeds into the legal financial system or, at the opposite, on the value added produced by criminal activities at large. While the former approach is complex due to the great effort required for the meticulous identification of predicate offences and related illicit financial flows introduced in the legal system, widening the definition of reference or adopting a general macroeconomic indirect approach to the issue will introduce also further complexities.

Other factors, such as considering self-money laundering, the nationality of ML subjects and banks involved, can yield diverging results with varying levels of accuracy. Identifying the correct country of ascription of ML is crucial to distinguish between 'imported' and 'exported' amount. Failure to do so could lead to underestimates, overestimates, or double counting of ML for the country under examination.[11]

Many authors criticize the estimation methods used in the empirical literature due to questionable hypotheses, approximations, and extrapolation methods (Walker and Unger, 2009; Reuter, 2013; Unger, 2013, Walker and Unger, 2013; Levi et al., 2018). They often point out examples that are widely cited as "facts by repetition". One of the most emblematic instances is the citation of ML estimates provided by the Managing Director of the IMF, Michael Camdessus, during a speech at the plenary meeting of the Financial Action Task Force on ML in 1998: "*(…) While we cannot guarantee the accuracy of our figures – and you have certainly a better evaluation than us – the estimates of the present scale of money laundering transactions are almost beyond imagination – 2 to 5 percent of global GDP would probably be a consensus range.*". These numbers were later reinforced by a report of UNODC (2011), which estimated ML related to drug trafficking at 1.6 trillion dollars (2.7% of world GDP) and that referred to all predicative offences for ML at 2.1 trillion dollars (3.6% of world GDP). Placing this percentage in the middle of the IMF range ended in emphasizing the significance of both estimates.

Other scientific studies provide ML figures that, although developed on the basis of theoretical frameworks and diverse empirical methods, significantly differ from the above mentioned percentages.

First, there are studies following the well-established Currency Demand Approach (CDA), which analyse the demand for currency in circulation by comparing it with sources of legitimate economic transactions and identify excess cash flows that may be indicative of ML.[12] Among these, Ardizzi et al. (2014) propose a revised CDA approach. By estimating an econometric micro-founded model for Italy over the period 2005-08, they break down cash inflows credited to current banking and postal accounts into legal, shadow and ML components, quantifying the latter between 6% and 8% of national GDP. In a previous scientific study on Italy, Zizza (2002) estimated the size and evolution of the shadow economy, also considering the impact of criminal activities on the demand for cash payments as a variable in a CDA model. The findings were consistent with the official figures provided by ISTAT for the same years, once the demand for cash linked to criminal activities was excluded. More recently, Aljassmi et al. (2024) estimate the magnitude of ML in the UAE using a similar approach. Their results illustrate that an amount equivalent to about 19% of the GDP has been laundered on average in that country between 1975 and 2020.

Another branch of the literature relies on the development of theoretical models and simulation techniques. These models typically incorporate data on various aspects of the economy, financial system, and criminal behaviour to simulate the flow of illicit funds. By calibrating the models with available data

[10] Relevant examples include the cases of undeclared work – which is considered as predicate offence in United States while it is not in Germany and in the Netherlands – or light drugs, such as hashish and marijuana, or prostitution which are legal in the Netherlands and illegal in many other countries (Unger, 2013).

[11] Moiseienko and Keatinge (2019), for example, formalize a breakdown of national ML into the following three components: domestic laundering of domestic illegal proceeds; 2) national laundering of proceeds generated abroad; 3) laundering of proceeds generated abroad through the involvement of national business infrastructures.

[12] Original works with application of CDA to underground economy and ML are those of Tanzi (1980; 1983; 1996).

and adjusting key parameters, researchers can generate numerical estimates of the magnitude of ML. Argentiero et al. (2008) calibrate a dynamic general equilibrium model using observable and estimated macroeconomic data, finding that ML in Italy accounts, at least, for 12 percent of aggregate GDP. Barone and Masciandaro (2011), for example, propose a theoretical model to analyse the effects of the reinvestment of laundered money into legal markets. They find that the legal reinvestments of criminal organizations increase when illegal returns rise, the share of illegal revenues to be laundered decreases, initial illegal revenues are higher and the index of ML regulation laxity is higher.[13] Moreover, they observe a higher volatility of ML relative to GDP and a negative correlation between the two. Astarita et al. (2018) develop a theoretical model to evaluate the effect of four main crimes attributed to criminal organizations: extortion, corruption of public officials, trade in illegal goods and ML. Despite the ex-ante undetermined effect of criminal organizations on economic activity,[14] their simulation exercise delivers a negative effect on the level of the economic activity and the growth process for Italy, with a key role played by the trade in criminal goods and ML. Lastly, Loayza et al. (2019) estimate the size of illicit income derived from cocaine exports and common crimes and provide simulated and econometric estimates of asset laundering in the Colombian economy; they find that illicit incomes and ML peaked at nearly 12% of GDP in the 1990s and then decreased to less than 2% by 2013.

Other papers adopt macroeconomic approaches and various types of econometric estimations. Walker and Unger (2009) propose a revised version of the original gravity model of Walker (1995) and provide estimates of global flows of illicit finance decomposed by crimes, accounting for a range of 1,061-1,599 US billion dollars. However, the robustness of such approaches is negatively affected by the set of hypotheses and methodological simplification required in order to derive the estimation. Among the most important, it is enough to cite the need of a baseline estimation of criminal proceeds by type of crimes, the corresponding hypothesis about the prevalence of different types of crime and the average proceeds per crime, the proportion of those proceeds that are likely to be laundered and the estimates of the proceeds from foreign crime that flow into a country for being laundered. Schneider (2010) uses a latent estimation procedure (MIMIC: Multiple-Indicators Multiple-Causes) and shows that the turnover of organized crime activities increased from $270 billion in 1995 to $614 billion in 2006 for 20 OECD countries. Other articles seem more focused on providing a multitude of figures that are only marginally relevant to understanding the phenomenon. These numbers often appear only weakly connected to each other, possibly cited simply to emphasize the complexity of the issue and raise awareness about it (Schneider and Windischbauer, 2008).

Ferwerda et al. (2020) use – for the first time to our knowledge – a dataset of financial transactions suspicious of ML provided by the Dutch Institute infobox Criminal and Unexplained Wealth (iCOV)[15] to simulate all ML flows worldwide, distinguishing between laundering of domestic crime proceeds, international investment of dirty money, and money just flowing through each country. At the first step a gravity model for ML (the dependent variable) is estimated[16] on a set of largely used explanatory variables,[17] using data on STRs aggregated at level of year and foreign country from which financial flows

[13] To this last respect, since the cost of ML depends on the effectiveness of the AML regulation, every improvement in the effectiveness of AML regulations would produce a decrease in the value of ML activity, which corresponds to an increase in overall public welfare.

[14] Indeed, on one hand it decreases aggregate demand by draining resources through extortion, corruption and the implied consumption of criminal goods, on the other hand it increases demand through ML when the ill-gotten money is used for the consumption of legal goods and re-investments in legal activities.

[15] In the Netherlands, obliged entities are asked to send the so-called Unusual Transaction Reports (UTRs) to their FIU; then these are filtered in order to select the reports to be passed on to the law enforcement agencies (i.e. which UTRs are STRs).

[16] Perhaps, this is the greatest innovative contribution of Ferwerda et al. (2020), since previous studies using 'gravity models' (e.g. Walker, 1999; Walker and Unger, 2009) actually did not estimate it. In fact, since the value of ML was unknown, they simply adopted 'qualitative hypotheses' to identify the coefficients to apply to each determinant of ML and, only by derivation, obtained an arithmetical evaluation of its amount.

[17] These include, among others, GDP (the usual 'mass variable'), GDP per capita, distance, corruption and 'tax havens' indicators, use of a common currency, common religion and language, border dummy and annual value of bilateral trade.

enter the Netherlands and vice versa. Then, on the basis of several hypotheses, they simulate the international flows between 187 countries[18] and show that ML generally amounts to 1.9% of GDP for OECD countries and 3% in the World average.[19] For what concerns Italy, they estimate total ML being 1.3% of GDP.

For what concerns Italy, Pinotti (2015) estimates the economic impact of organized crime in the regions of Puglia and Basilicata in the early 1970s, resulting in a GDP loss of about 16% over three decades (about 0.5% on an annual average). Using a similar econometric approach, Barone and Mocetti (2014) demonstrate how the inflow of public funds into earthquake-affected areas of Friuli and Irpinia (in 1976 and 1980, respectively) generated different effects in the two regions: over the following thirty years, in Friuli Venezia Giulia – characterized by negligible levels of organized crime – per capita GDP growth was about 20 percentage points higher than that observed in a counterfactual region, while in Irpinia – where organized crime was strongly rooted – per capita GDP growth was about 12 percentage points lower than that of the control region.[20]

Finally, as a complementary approach, other studies use qualitative information to gain insights into ML and related phenomena. Levi et al. (2018) analyse the National Risk Assessments[21] of five major countries, revealing the limited use of data in evaluating the AML systems. They acknowledge that, without reliable data analysis, the evaluation of countries' AML effectiveness "*will be open to allegations of ad hoc, impressionistic or politicized judgments*" (p. 307). At the same time, they also admit that estimating the size of a phenomenon can be a useful tool for prioritizing public resources. The problem of unreliable estimations arises from many factors, including the lack of relevant data, robust methodologies, and inherent theoretical difficulties. These challenges persist despite international efforts to definite standards and common principles.

Other types of approaches – not explicitly included in this review – include the use of case studies to infer general estimates, other qualitative analysis, and direct estimation methods, such as surveys, panel interviews or analysis of data on assets confiscated by law enforcement agencies.

4. A data-driven approach for measuring Money Laundering in Italy: methodology and data

4.1 Methodology

Statistics is a missing data problem and the goal is to predict unknowns with appropriate measures of uncertainty (Little, 2013).

The core idea underlying the estimation method which is proposed in this study is drawn from the context of data imputation in sample surveys. Indeed, one can conceptualize the STRs reporting process as a general data collection process wherein the statistical units (or 'respondents') may furnish erroneous information, overlook it completely, misunderstand a certain question, or be unable to provide the correct answer to a certain question, either unintentionally or deliberately. In such cases, the resultant database of the collection phase might appear accurate for some units (i.e. exhibiting non-missing and true values

[18] Since the only real data used come from the Netherlands, the authors admit that is more likely that the results hold for rich and developed countries than for poor and underdeveloped countries, that is, only for 36 OECD countries.

[19] Relatively small countries seem to be used mostly as through-flow countries, while other countries (Germany and US included) are on the other end of the spectrum with the main ML challenge being the laundering of domestic crime proceeds. Most ML should happen in the United States and the United Kingdom (40% of all ML in the 36 OECD countries), while as a percentage of GDP ML is highest in Belgium, Luxembourg, and Israel and lowest in Japan and South Korea.

[20] This is mainly due to the degree of corruption which has led to a distorted allocation of resources and, consequently, to a decrease in the economic efficiency of the local economic system.

[21] Many countries conduct National Risk Assessments (NRAs) to evaluate the threats and vulnerabilities related to ML within their jurisdictions. NRAs involve a comprehensive analysis of various factors, including the legal and regulatory framework, financial sector integrity, law enforcement capabilities and international cooperation. While NRAs may not directly estimate the size of ML, they provide valuable insights into the factors driving illicit financial flows.

for each variable) while affected by errors for others, that is with missing or erroneous data for one or more variables of interest.

In broad terms, when confronting certain measurement errors or missing data for a quantitative attribute, there are two main alternative strategies to adopt: unit removal or data imputation. The former approach entails 'listwise deletion', where only units with reliable and complete data are retained throughout all analytical stages, or 'pairwise deletion', where summary statistics are computed from time to time only for units with available information. On the contrary, the latter involves maintaining the number of units unchanged and imputing missing data with plausible values.

Without going into methodological issues and analytical consequences of adopting one approach over the other, let us to contextualize the aforementioned framework within the scope of this paper. We consider a conventional aggregate unit of reference of the STRs reporting system – namely, the combination of bank and province identifiers – as affected by measurement error, also including the extreme case of missing data, for the quantitative attribute 'value of money laundering'. Hence, our dataset comprises bank-province combinations for which we have complete and accurate information regarding contextual data (as delineated in Table 1) while only a subset of them exhibits comprehensive and reliable information on financial transactions associated to ML.

It is important to note that the distribution of financial transactions suspected of ML across bank-provinces tends to vary significantly from year to year, with changes also occurring suddenly without the observation of evident predictive factors. A bank in a particular territory, which may have not been involved in ML cases in the past, could suddenly experience a significant episode of this nature. It follows that a longitudinal approach could yield estimates that fail to adapt to the evolving nature of the phenomenon and relying on historical data to predict current or future trends may result in outdated or inaccurate assessments.

The choice of examining banking operations at the provincial level, rather than at a different geographic unit, is based on its effectiveness in capturing relevant factors associated with ML. This level of analysis provides a balanced approach for contextualizing local financial dynamics. Using regions as the units of analysis might dilute the precision of the model estimates regarding the relationship between variables related to ML. On the other hand, focusing on smaller units, like municipalities, could introduce biases due to undetectable correlations between neighbouring areas. Provinces serve as a practical and effective territorial unit where threats and vulnerabilities for the financial system can be more accurately connected with local criminal networks or broader illegal behaviours. Additionally, focusing on the financial characteristics of banking institutions at the provincial level is also acceptable since it likely aligns with the business and organizational models that banks adopt in order to better match clients' demand of financial services emerging at local level. When considered alongside financial and socio-economic frameworks, these characteristics can shed light on potential opportunities for the exploitation of financial institutions by money launderers.

In the rest of this section we describe in detail the adopted estimation process. It can be divided in the following three main phases: 1) selection of relevant ML reports; 2) identification of unreliable aggregate values of ML reports; 3) imputation of unreliable aggregate values of ML reports and resultant final estimate.

Selection of relevant money laundering reports

As a fundamental first step, we employ a rigorous selection of relevant STRs to build a database of 'observed ML' aiming to closely align it with the phenomenon under analysis. This phase primarily addresses one of the main critique raised by several experts in the field regarding the use of AML reports to evaluate the scale of ML.[22] This critique arises from the view that STRs do not necessarily capture such unlawful activities only, as some of them may be proved, upon investigation, to be legitimate transactions. To address this concern, we consider only reports with financial and investigative relevance, that is, those

[22] See, for example, Walker (1995) and Reuter (2013).

which, following the financial analyses conducted by UIF experts, either received a final risk score of 4 or 5 (on a scale ranging from 1 to 5) or positive investigative interest from law enforcement agencies.[23]

This approach enables us to establish a database based on strong suspicions of ML; while it does not perfectly align with the most rigours available definition of ML, instead it serves as the best available proxy for the phenomenon under investigation, albeit with inherent biases. Note that everything that follows, actually the core of the proposed methodology, is basically devoted to the correction of such lack of perfect knowledge. The importance of the selection phase is evident by recognizing that banks have to monitor financial transactions and classify them as suspicious (by following international standards and the Italian law): the actual relevance in terms of ML can only be achieved after financial analysis and investigative feedback.

Other general selection criteria have been applied by including : *i*) STRs whose financial analysis was actually concluded (i.e. STRs sent to the competent Enforcement Agency); *ii*) financial transactions actually labelled as 'suspicious';[24] *iii*) financial transactions that were executed (not only attempted) by bank clients; *iv*) STRs not archived by the UIF (i.e., for which sufficient risk elements were found to support the suspicion of ML); *v*) financial transactions that are fully detailed and properly listed in the STRs. Every obliged entity is requested to report also the 'overall amount of suspicious financial operation' without providing the UIF with the necessary auxiliary information to assess its accuracy.

As the final step in data selection, we identify and delete duplicate financial transactions sent by the same bank in order to avoid or significantly reduce double-counting. Note that including the same financial transaction in different STRs is fully legitimate, as it may be flagged as suspicious initially and then reappear as relevant as further suspicious financial activities of the same person or legal entity are investigated.

Identification of unreliable aggregate values of money laundering reports

The second phase of the methodology also addresses another key criticism regarding the use of STRs, namely that the reported financial transactions would usually represent only a partial amount of a larger chunk. To this end, we distinguish between consistent and inconsistent reports of suspected ML by employing a machine learning procedure to estimate the monetary value of the financial transactions suspected of ML that each bank-province was expected to report based on a set of control variables. Specifically, we utilize the Quantile Random Forest (QRF) algorithm, which is described in detail below, to estimate every percentile of the observed annual distributions of such subset of financial transactions. This allows us to identify quantile-based thresholds, enabling the distinction between bank-province observations featuring a size of ML roughly in line with the expected value and those that are inconsistent with it.

As a result of this, we can categorize the entire sample of bank-provinces into the following three groups:

1. Under-reporting bank-provinces, characterized by an aggregate value of reported ML below a lower (or 'left-side') threshold;
2. Over-reporting bank-provinces, exhibiting a reported value of ML above a higher (or 'right side') threshold;
3. Reliable bank-provinces, with a reported value of ML falling between the two thresholds.

The first two groups of observations encompass cases commonly identified as problematic when utilising AML reports for the estimation of ML size. An STR may represent merely a fraction of a larger illicit operation (i.e., it undervalues the actual financial value of ML) or may involve financial transactions entirely legitimate (i.e., it overvalues the actual financial value of ML). However, when a significant

[23] In particular, we consider the most recent investigative interest available at the date of data extraction.

[24] When compiling an STRs, banks can also report non-suspicious financial transactions which, however, can be useful to the UIF's analysts for contextualising the suspicious ones.

portion of suspicious operations is consistently identified by a large share of other bank-provinces (group 3), we can acknowledge them as statistically reliable and use the associated amount of financial operation to correct the under-valued or over-valued estimates from the first two groups, providing a more accurate value.[25]

It is worth noting that bank-provinces lacking valid ML financial information or not reporting STRs at all do not necessarily coincide with those requiring data imputation. Instances of zero ML occurrences might be considered acceptable when assessed against a set of context variables (hereafter referred to as 'controls'). Conversely, even substantial financial transactions classified as high-risk for ML may be considered unreliable if they are inconsistent with those controls.

As mentioned at the beginning in this section, we opted to treat the multi-year information as separate cross-sectional datasets for both the categorization of bank-provinces and the subsequent phase of data imputation. This approach allows us to avoid the use of time series or panel data methods, which might be particularly risky in the context of ML. Indeed, while time-series and panel analysis are usually effective for analysing household consumption or firms' investment decisions, the unique nature of ML complicates their application. Unlike more commonly studied economic behaviours, ML activities do not consistently occur in the same location (whether a specific bank or geographic area) with predictable intensity or regularity.

Let us formalize the above framework and denote the target variable as Y (the value of ML), X as the set of covariates, and $F(y/X = x) = P(Y \leq y | X = x)$ as the conditional distribution function of Y, that is the probability that Y is smaller than y, for a given $X = x$. It follows that the general α-quantile of the distribution of Y is defined by $Q_\alpha(x) = \inf \{y : F(y/X = x) \geq \alpha\}$: the probability that Y is smaller than $Q_\alpha(x)$ for a given $X = x$ is exactly equal to α.

In general, quantile regressions aim to estimate the conditional quantiles from data. While various parametric and non-parametric approaches can be utilized for empirical estimation (Koenker, 2005; Koenker et al., 1994; Chaudhuri and Loh, 2002), the *QRF* algorithm introduced by Meinshausen (2006) offers a method based on random forests (Breiman, 2001) that is particularly suitable for our case. Unlike standard random forests, which approximate the conditional mean of the response, *QRF* models the entire conditional distribution of the response variable given a set of explanatories X using a random forest of regression trees.[26] This allows for the estimation of all the quantiles of the conditional distribution of Y/X, enabling the identification of potential 'outliers' falling outside the general prediction interval $[Q_\alpha(x), Q_{1-\alpha}(x)]$.

While the choice of *QRF* for identifying unreliable (i.e., anomalous) bank-provinces is only one option among many possible supervised and unsupervised approaches, it stands out for several desirable properties. These include its ability to leverage relevant auxiliary information pertinent to the phenomenon under investigation and its capacity to define thresholds based on easily understandable concepts such as distribution quantiles.[27] Beside these general aspects, *QRF* possesses several advantages over other approaches such as standard quantile regression (*QR*) or Quantile Gradient Boosting (*QGB*):

1. *QRF* provides simultaneous estimation of each quantile of Y/X, unlike *QR* and *QGB* which require different models or algorithms for each quantile;

[25] Gara and Pauselli (2020) discuss measurement issues in banks' reporting of suspicious transactions to supervisory authorities. They underline that banks may have an incentive to over-report STRs due to the asymmetric treatment of Type I and Type II errors (as sanctions apply only to omitted reports). However, in this phase we care of both possibilities.

[26] This is because, while RF considers only the mean of the observations of Y falling into each node and neglects all other information, QRF preserves all the information (the value of each observation) thus evaluating the entire Y/X conditional distribution.

[27] Lenza et al. (2023) apply the same algorithm to the median forecasts of euro area inflation, while Cusano et al. (2022) use it for identifying errors in banks' supervisory reports on loans to the private sector employed in the Bank of Italy's statistical production of Monetary and Financial Institutions' (MFI) Balance Sheet Items (BSI).

2. The previous property implies that – differently from *QR* and *QGB* – *QRF* ensures no quantile crossing, meaning that the estimated value of a given quantile of *Y* cannot be greater (less) than the estimate of a next (previous) quantile;
3. The number of grown trees in *QRF*, typically assumed between 300 and 500, is not a significant adjustment parameter (i.e., the final result is commonly robust against it);
4. *QRF* yields robust results also with respect to the choice of the minimum number of observations in each terminal node, typically set at 10;[28]
5. While *QRF* may be weaker than *QGB* at predicting central quantiles, it offers more accurate prediction of extreme quantiles,[29] which is valuable to our scope of identifying anomalous observations.

Last but not least, the full non-parametric nature of *QRF* allows for avoiding any questionable hypothesis regarding variable transformations (logarithmic or arguable standardizations) as well as the functional relationship between control variables and the target variable.

However, the only partial limitation connected to the idea of using distributional quantiles as thresholds of data reliability relies on an *a priori* hypothesis about the expected proportion of observations to be considered as reliable/unreliable. To address this controversial issue, we adopt a conservative approach by providing a prediction interval instead of a point estimate. This is derived from multiple assumptions about the share of lowest and highest unreliable observations, ranging from the 5^{th} to the 15^{th} percentiles and from the 85^{th} to the 95^{th} percentiles of each annual distribution of ML values. This is equivalent to speculate the share of bank-provinces reporting inconsistent values of ML between 10% and 30%. It is worth noting that replicating the analysis across different thresholds also serves additional purposes: it mitigates the potential impact of misspecification of the target variable (not perfectly corresponding to a formal definition of ML).

The only assumption implicitly introduced by our choice is the symmetric distribution of the lower and upper thresholds, which could possibly be relaxed, with an equally questionable asymmetric choice, to better align with the uneven distribution of the value of anomalous aggregated reports.[30]

Imputation of unreliable aggregate values of money laundering reports

As the third and final phase, we utilize a version of the Predictive Mean Matching (*PMM*) method, as developed by Little (1988) on the basis of the fundamental contribution to statistical matching of Rubin (1986), to impute unreliable values of ML reports.[31] This involves replacing the aggregate value of financial transactions reported by each unreliable bank-province (under- or over-reporting) with the value observed for the most statistically similar observation selected among the set of reliable ones.

Formally, let I and J represent the sets of reliable and unreliable bank-provinces (i and j, respectively), and let $\boldsymbol{X} = (X_1, \dots, X_q)$ denote the matrix of control variables of Table 1[32] observed for both the units of I and J. The value of financial transactions suspected of ML recorded for each j (Yj) is replaced by the value observed for the unit i (Yi) that minimizes the following distance function:

[28] See Meinshausen (2006).

[29] See Yuan (2015).

[30] As many socio-economic phenomena, the distribution of the value of suspicious operations is indeed positive and right-skewed.

[31] Little (1988) proposed to apply a predictive mean neighborhood method with the use of a Mahalanobis metric defined on the basis of the regression residuals, while standard PMM imputes missing data randomly selecting a donor from a restricted set of non-missing units (with similar predicted values) and uses the corresponding observed value to replace the missing one (Van Buuren, 2018).

[32] For what concerns the indicators derived from the SARA database, only the amounts in euros have been included among the regression covariates. Including the number of financial transactions for each indicator alongside the amounts does not significantly alter the final estimates.

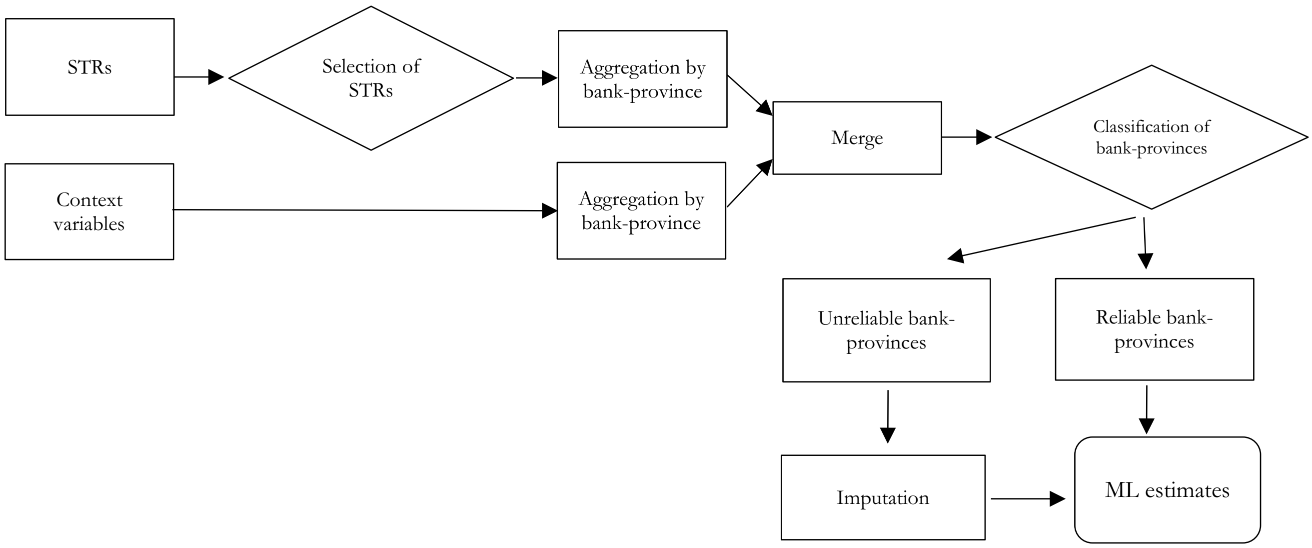
$$d^2(i,j) = \left(\hat{\mu}_i - \hat{\mu}_j\right)^T \hat{R}^{-1}_{Y \cdot X/I}\left(\hat{\mu}_i - \hat{\mu}_j\right)$$

where $\hat{\mu}_k = \hat{\mu}(X_k)$ represents the predicted mean for the generic unit *k* obtained from the OLS regression of $\boldsymbol{Y}$ on $\boldsymbol{X}$ estimated by using only the reliable bank-provinces of the set *I*, and $\hat{R}_{Y \cdot X/I}$ denotes the residual covariance matrix of the same regression. As specified by Little (1988), the usage of the elements of $\hat{R}_{Y \cdot X/I}$ as correction factors of the standard Mahalanobis distance can be motivated by observing that "*from a statistical stand-point, one might tolerate greater matching error for Y variables that are subject to greater prediction error*" (p. 291).

We perform this third phase separately for each of the twenty Italian regions and repeat it for eleven hypotheses regarding the share of bank-provinces to be identified as anomalous in the second step (ranging from 5% to 15% on either tail of the distribution). The first choice allows to derive an overall country estimate that by construction fully aligns with the regional-level estimates; the second one provides a more realistic interval estimate of the size of national ML, rather than a single point estimate.[33] Note also that this requires to run 220 regressions for each year. Thus, in order to maintain the number of significant regressors reasonably low (and consistent) and to prevent overfitting of a single complete model, a backward selection of context variables has been employed in each regression. The potential negative effect on the estimates accuracy is limited by the well-known *PMM* property of providing results robust to alternative specification of the underlying regression model.

Figure 1 presents a flowchart that outlines all the key steps of the proposed methodology. This visual representation systematically details each phase of the process, providing a clear overview of the methodology's structure and flow.

Figure 1. Key steps of the methodology employed for measuring ML in Italy



4.2 Data

The empirical application of this paper mainly relies on the database of Suspicious Transaction Reports (STRs) received by the Financial Intelligence Unit for Italy (UIF) during the period 2018-2022. The UIF is institutionally in charge to receive and analyse STRs, submitted by financial intermediaries, professionals and other qualified operators, with the aim of contrasting and preventing ML, financing of terrorism and proliferation of weapons of mass. These entities are mandated to identify, assess and promptly report such suspect transactions under the obligation of active cooperation. The legislative

[33] Differently from a point estimate – a single value estimate of a distribution parameter – an interval estimate provides a range of values where the parameter is likely expected to lie. The latter should not be confused with the concept of confidence interval, instead consisting of an upper and a lower bound of a point estimate one can expect to obtain, given a certain level of confidence, by resampling the population.

framework empowers the UIF to conduct uniform and integrated evaluations of the reports by capturing subjective and objective connections and networking dynamics, tracing financial flows transcending Italy's borders through information exchanges with foreign FIUs. Upon completion of each financial analysis, the reports are sent to the NSPV – the Special Currency Police Unit (*Nucleo Speciale di Polizia Valutaria*) and the DIA – the Antimafia Investigation Department (*Direzione Investigativa Antimafia*) for investigative follow-up. UIF also dispatches reports and analyses to the judicial authorities in the presence of crime-related information or upon their direct request; findings may also be shared with supervisory authorities if significant profiles are identified. Finally, data and information may also be exchanged with the DNA – the Antimafia National Directorate (*Direzione Nazionale Antimafia*) to scrutinize potential links between suspicious transaction reports and organized crime contexts, and enable prompt subsequent actions.

The choice to use STRs to estimate the value of ML is motivated by the fact that, despite its imperfections, not a better alternative exists for data regarding ML activities. To support this opinion, it is worth noting that upon receipt of a new STR, its content is cross-referenced with diverse data sources of internal or external origin and, notably, with historical archives of previous STRs. With the aid of a technological system developed by the UIF to manage this vast amount of data, supplemented by information from other public and confidential sources as well as from national authorities and foreign FIUs, the identification of authentic instances of ML is increasingly becoming a more robust and effective process. However, while under-reporting of relevant STRs and the inclusion of irrelevant information may still occur, the estimation methodology proposed is expected to largely correct for these issues, with such cases likely remaining minimal and showing a decreasing trend over time.

In 2022, the last year the following analysis refers to, UIF received 155,426 STRs related to hundreds of thousands of individual financial transactions and entities (individuals or legal units). The majority of reports were submitted by banks (57.3%), followed by non-bank financial intermediaries (30.2%), gaming service providers (6.0%), professionals (3.6%) and non-financial operators (2.1%).[34] For the purpose of our analysis, we exclusively consider STRs filed by banks. Notably, transactions reported by banks in 2022 accounted for 80.5% of the total monetary value of all STRs, followed by professionals (8.9%) and non-bank financial intermediaries (8.5%).[35] Given the increasing share of suspicious financial transactions reported by non-bank financial intermediaries, it is realistic to expect that future developments of the applied approach will likely incorporate these entities. On the contrary, professionals, gaming service providers, public administration and other non-financial operators will continue to be excluded from the analysis as long as the current methodology is in use, due to the lack of data concerning financial variables – other than STRs – used along the main phases of this study.[36]

For the scope of this paper, we decided to aggregate all the financial information available at the UIF by bank and province where the transactions took place. The complete set of variables employed to proxy the economic and financial fundamentals affecting the level of ML is summarised in the following Table 1. The first part of the table delineates the variables of primary interest for the estimation procedure and several composite indicators derived from the Aggregate Anti-Money Laundering Reports (S.AR.A. from the Italian acronym). In accordance with Italian AML legislation (Legislative Decree no. 231/2007), banks and other financial intermediaries have to report, irrespective of any suspicion of ML, all transactions amounting to € 5,000 or more on a monthly basis to the UIF, after aggregating them

[34] General government entities sent a residual 0.1% of reports. For a detailed picture on the number and share of STRs by type of reporting entity refer to Table 1.2 of the 2022 Annual Report of the Italy's Financial Intelligence Unit (UIF, 2023a).

[35] The complete data can be extracted by the table a.2.1 of the semiannual reports "Quaderni dell'antiriciclaggio dell'Unità di Informazione Finanziaria, I e II semestre" (UIF, 2022; UIF, 2023b. Available only in Italian).

[36] It should be noted that the exclusion of non-financial operators is likely to have only a marginal effect on our estimates. The amounts of money laundered through these entities is possibly very low if compared to that faced by larger financial actors and, more importantly, it would likely pass anyhow through bank accounts at some point.

according to several criteria.[37] Consequently, these reports represent one of the most exhaustive and detailed proxy of financial operations.

Table 1. List of variables.

Variables	Level	Source
Value of financial transactions included in STRs	bank-province	UIF
Value and number of total transactions	bank-province	SARA (UIF)
Value and number of cash transactions[1]	bank-province	SARA (UIF)
Value and number of transfers (distinctly in debit/credit) with countries at risk[1]	bank-province	SARA (UIF)
Value and number of off-account operations[1]	bank-province	SARA (UIF)
Value and number of unpaid cheques[1]	bank-province	SARA (UIF)
Value and number of banknote denomination exchanges[1]	bank-province	SARA (UIF)
AML bank classification[2]	bank	UIF
Incidence of crimes[3]	province	Istat
Shadow added value[4]	province	Istat
Taxable income per capita	province	Agenzia delle Entrate
Share of firms potentially linked to organized crime over total number of registered firms[5]	province	UIF

[1] Financial variable previously used by the UIF and the Bank of Italy to compute risk indicators at the bank level, originally developed by the AML supervisory unit of the Bank of Italy (see UIF (2017), pp. 83-84).
[2] Between 2018 and 2021 it distinguishes between: 1. Large banks; 2. Large banks respectively in macro-geographical area; 3. Large banks with specialized operations and foreign banks with traditional activities; 4. Small banks (in the macro geographic areas and abroad). For 2022, the following new classification is used: 1. Large banks; 2. Medium and small banks; 3. Private banking; 4. Corporate banks and investment banking.
[3] It includes the following crimes reported by the police force to the judicial authority (per 100,000 inhabitants): prostitution, robberies, extortion, counterfeiting, receiving stolen goods, smuggling, money laundering, usury, drug trafficking, criminal conspiracy, and mafia association.
[4] It includes the share of under-declaration and the share of added value attributable to irregular work (source: ISTAT).
[5] This indicator is derived using the methodology described in Cariello et al. (2024).

The SARA reports refer to the operations carried out by the customers of the intermediaries and, as the data are aggregated, they are filed in an anonymous format. Although SARA data are available on a monthly basis, we aggregate them annually, consistently with all other sources, and with the objective to provide yearly-based estimates of ML. Moreover, in order to align with the bank-province framework of the STRs database, only reports submitted by banks at the local level are used, also in this case excluding reports of non-bank financial intermediaries.[38] Specifically, we consider the total number and nominal value of all financial transactions, along with the number and value of operations referring to specific payment instruments, such as cash transactions, wire transfers between Italy and country at risk of ML,[39] off-account financial operations, unpaid cheques, and exchanges of banknotes by denomination. It is noteworthy that, aside from the aggregate total financial transactions which reflect the overall

[37] Aggregation criteria include the type of transaction, the intermediary's branch where the transaction took place, the client's residence (at municipality level) and his/her economic sector. Each aggregate record includes information on the total amount transacted, the corresponding cash component and the number of underlying individual transactions being aggregated.
[38] For a complete list of the categories of reporting entities see Table 5.2 of the 2022 Annual Report of the Italy's Financial Intelligence Unit (2023).
[39] The list of countries is updated yearly on the basis of official national and international documents. In 2022, for example, the list of non-cooperative jurisdictions and/or tax havens is drawn from the ministerial decrees implementing the Consolidated Law on Income Tax (TUIR), from the lists 'High-Risk Jurisdictions subject to a Call for Action (black list)' and 'Jurisdictions under Increased Monitoring (grey list)' published by the FATF in March 2022, from the 'EU list of non-cooperative jurisdictions for tax purposes' (February 2022 update) and from the list of countries identified by the European Commission with Delegated Regulation EU/2016/1675 and subsequent amendments.

financial size of bank operations, all other indicators are in some way correlated with an increased risk of ML.

Additionally, we incorporate several structural factors, including a four-item classification of banks based on size and geographical location, the per capita taxable income at the provincial level as a proxy for wealth, and two variables accounting for additional ML risks: the incidence of crimes and the extent of shadow economy at the provincial level, both sourced from Istat. Furthermore, we include a specific provincial indicator of economic infiltration, measured by the share of firms potentially connected to organized crime relative to the total number of registered firms in the province.

The connection between the STRs database and the SARA archive at the bank-province level may be marginally affected by financial transactions conducted by clients of online banks, because of uncertainty on the province where the operation took place registered in both the databases. Currently, this issue has been addressed as follows: for SARA records, a correction has been applied only to digital banks, which operate exclusively online, without physical branches. In such cases SARA operations are typically based on a single province, usually the province of the legal address. In order to capture more accurately the actual location of these operations, the customer's province of residence has been used as a proxy for the province of operations. Conversely, no correction has been applied, for the same banks, for the corresponding observations in the STRs database. Indeed, actual STR data have been showing, along the entire period of analysis, a wide territorial distribution of operations, presumably corresponding to the customer's actual place of financial operation or, at least, registered residence.

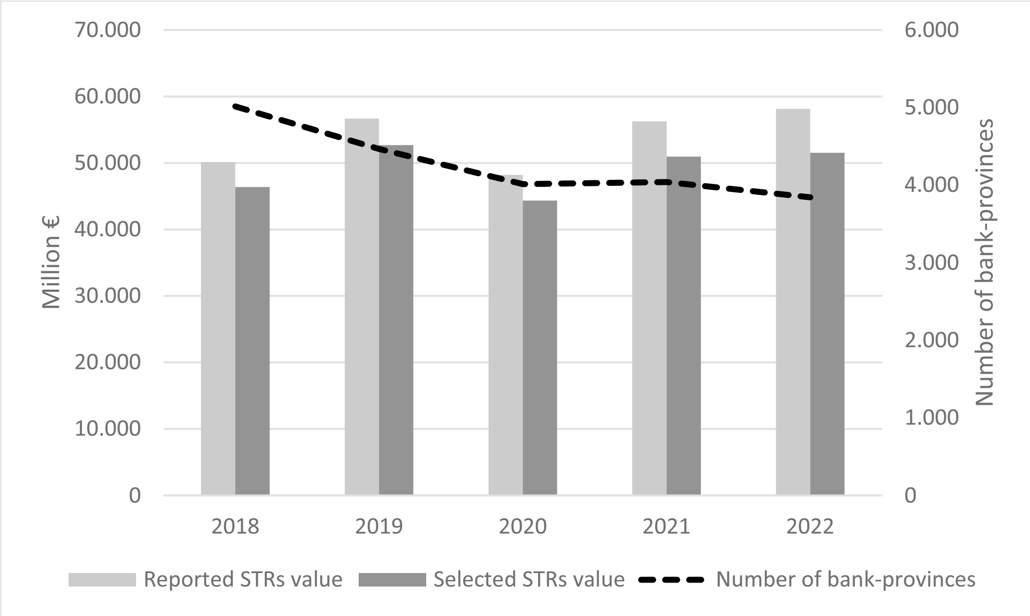
Overall, the proposed approach allows to provide an estimate of the size of ML by avoiding several methodological issues that other approaches fail to resolve. It does not rely on data indirectly connected with the definition of the phenomenon, such as data from predicate offences, preliminary estimates of the revenues (or net proceeds) of money launderers, etc. Moreover, it bypasses the need to explicitly measure exported and imported ML funds (i.e., financial resources ill-gotten within/outside the reference country but laundered outside/within it). What is recognised as suspicious by the financial system and reported to the Financial Intelligence Unit is the only thing that matter, regardless of the nationality of the launderer or the source of illegal assets and income. The most significant improvement concerns the phenomenon being estimated. For example, the indirect method reported in the introduction section relies on macroeconomic relationships among economic factors and indicators, which can be formalized by statistical models. However, it is unclear what they actually estimate. Is it the illegal component? Might it include irregular components or simply refer to the informal economy (not relevant to intercepting ML)? The use of AML reports, with due caution and appropriate methodological choices, allows overcoming these and other difficulties by relying on the most relevant available source of information, that is, detailed bank operations and the reporting of suspicious transactions at the local level.

5. Results

This section primarily aims to present the main result of the estimation of the value of ML in Italy. Figure 2 displays the total aggregate value of suspicious financial transactions reported by banks to UIF between 2018 and 2022. The figure distinguishes between two aggregate values. The first, labeled 'Reported STRs value', is obtained by applying general selection criteria. The second, referred to as 'Selected STRs value', is instead based on a targeted selection of STRs that exhibit strong indications of ML (high risk scores or significant investigative interest; selection described in Section 4.2).

Two main trends are evident in the figure: 1) the value of suspicious financial transactions selected on the basis of the criteria described in the previous section has been increasing over time (except for the peculiar year of 2020, influenced by the COVID-19 pandemic and the consequent economic and financial crisis), rising from 46.4 to 51.5 billion euros during the period 2018-2022 and representing, on average, 91% of the overall financial value reported; 2) the number of bank-provinces observed has been constantly decreasing due to the progressive closure of bank branches, a trend common in many economically developed countries, and decreased from 5,014 in 2018 to 3,843 in 2022.

Figure 2. Total value of financial transactions recorded in STRs (left axis) and number of bank-province combinations (right axis) - by year



Note: The total value of reported STRs does not consider: a) financial transactions that were attempted but not executed by bank clients; b) STRs received by the UIF but whose financial analysis was still in progress at the time of the analysis; and c) STRs sent by banks but archived by the UIF (i.e., for which sufficient risk elements were not found to support the suspicion of ML). The total value of selected STRs is derived from the reported ones by excluding cases of double-counting and considering only STRs that received a risk score of 4 or 5 (the highest two levels on a scale ranging from 1 to 5) or significant investigative interest from law enforcement agencies.

Note that since the estimation strategy mainly relies on the availability of context variables provided by the SARA database, the number of bank-provinces considered each year is driven by the occurrences observed in this data source.[40]

Table 2 presents key summary statistics concerning the control variables used in the second and third phases of the estimation procedure. First, it highlights the significant financial value recorded in the SARA database. For example, total financial transactions amounted to nearly 28 billion euros in 2022, corresponding to more than 14 times Italy's GDP. Although the SARA database tracks financial flows, which differ from the cross-section of goods and services produced (as provided by the national GDP), it underscores the relevance of this data source as an excellent proxy for local financial operations. In other words, the actual financial activities of banks across Italy are represented in a unique fashion by the information recorded in the SARA database. Other noteworthy patterns include the decreasing (nominal) value of cash transactions, which contrasts with the increasing trends of other types of financial operations (wire transfers, off-account operations, unpaid cheques, and banknote denomination exchanges).

[40] It is important to emphasize that banks not reporting SARA data are very few and typically have minimal operations. Nevertheless, these banks may still report STRs. The share of unmatched financial transactions reported in STRs varies between 0.3% and 0.8% during the five-year analysis period. While the exact impact on the final estimate of ML provided at the end of this section is unknown, it is reasonable to assume that it is very limited.

Table 2. Summary statistics of the context variables

	2018	2019	2020	2021	2022
Total financial transactions (millions of €)	18,959,736	19,562,872	18,883,076	22,814,235	27,961,337
Cash transactions (millions of €)	204,771	200,231	158,467	158,599	175,269
Inward wire transfers with countries at risk of ML (millions of €)	137,527	155,701	143,380	185,430	249,212
Outward wire transfers with countries at risk of ML (millions of €)	129,256	141,150	120,867	154,923	226,381
Off-account operations (millions of €)	1,424,868	1,290,465	1,267,670	1,484,751	2,200,554
Unpaid cheques (millions of €)	29,666	29,447	31,278	21,809	29,722
Banknote denomination exchanges (millions of €)	21.6	28.7	16.1	20.8	22.4
Shadow value added (millions of €)	170,493	164,422	139,532	157,431	157,431
Taxable income per capita provincial means (€)	19,492	19,690	19,605	20,458	21,433
Crimes (no. for 100.000 inhabitants; provincial means)	146	136	119	118	124
Provincial indicator of firms infiltrated by organized crime (provincial mean)			0.019		
AML bank classification					
bank type 1	1,905	1,632	1,383	1,382	1,405
bank type 2	1,200	939	762	764	1,885
bank type 3	1,194	1,198	1,193	1,204	479
bank type 4	715	697	674	690	74
Total number of bank-provinces	5,014	4,466	4,012	4,040	3,843

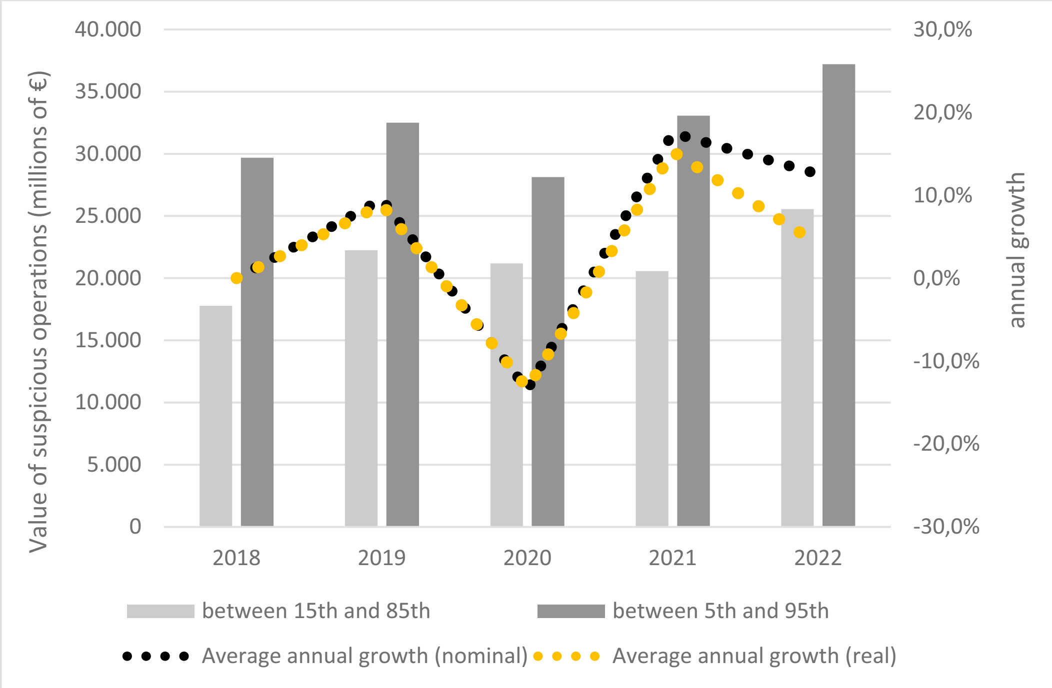
Note: the shadow value added of 2021 has been also used for the year of 2022 for shortage of updated data; for the period 2018-2021 the variable 'bank type' distinguishes between: 1. Large banks; 2. Large banks respectively in macro-geographical areas; 3. Large banks with specialized operations and foreign banks with traditional activities; 4. Small banks (in the macro geographical areas and abroad); for 2022, the following new classification is used: 1. Large banks; 2. Medium and small banks; 3. Private banking; 4. Corporate banks and investment banking. All euro (€) figures are expressed in nominal terms.

As discussed in Section 4, we proceed with the identification of unreliable and reliable aggregate values of ML reports at bank-province level by applying the *QRF* algorithm.[41]

Figure 3 shows the aggregate nominal value of the suspicious financial transactions reported by reliable bank-provinces based on two opposite hypotheses about their share over the total: a more conservative assumption (70%) and a wider share (90%). The figure highlights that, between 2018 and 2022, the total value of reliable transactions suspected of being connected to ML has been increasing, except for 2020, both in nominal and real terms.

[41] Other statistical techniques, commonly used for identifying potential outliers, have also been tested to detect unreliable bank-province units. These techniques include the interquartile range criteria, standard quantile regressions and its variations for modelling zero-inflated outcomes (Ling et al., 2022). Despite the theoretical appeal of these methods, none seems to have yielded results as robust and reliable as those obtained with the QRF approach. The properties of QRF, as previously described, sustain its use in addressing the complexities of the dataset and of the goal to achieve, reinforcing its suitability for this analysis.

Figure 3. Total value of suspicious operations reported by reliable bank-provinces, by year and two different hypotheses concerning the quantile threshold



Note: the dotted lines are obtained by averaging the value of financial transactions reported by reliable bank-provinces over the 15th and 85th quantile threshold and the 5th and 95th quantile threshold.

The empirical procedure described in Section 4 provides an estimate of the value of ML and other illicit conducts in Italy over the period 2018-2022 (Table 3), averaging 1.8% of GDP.[42] Putting the results in terms of interval estimation – according to the hypothesises concerning the expected share of reliable bank-provinces – ML is estimated between 1.5% (average of the lower bound) and 2.0% (average of the upper bound) of the national GDP.

Table 3. Estimates of money laundering

Year	ML estimate			Suspicious financial transactions observed in reported STRs	
	Mean bn. €	Mean % of GDP	Interval estimation % of GDP	bn. €	% of GDP
2018	28.7	1.6	1.3 - 1.8	46.4	2.6
2019	32.6	1.8	1.6 - 2.1	52.7	2.9
2020	28.6	1.7	1.6 - 1.9	44.3	2.7
2021	33.3	1.8	1.5 - 2.1	51.0	2.8
2022	35.6	1.8	1.7 - 2.1	51.5	2.6

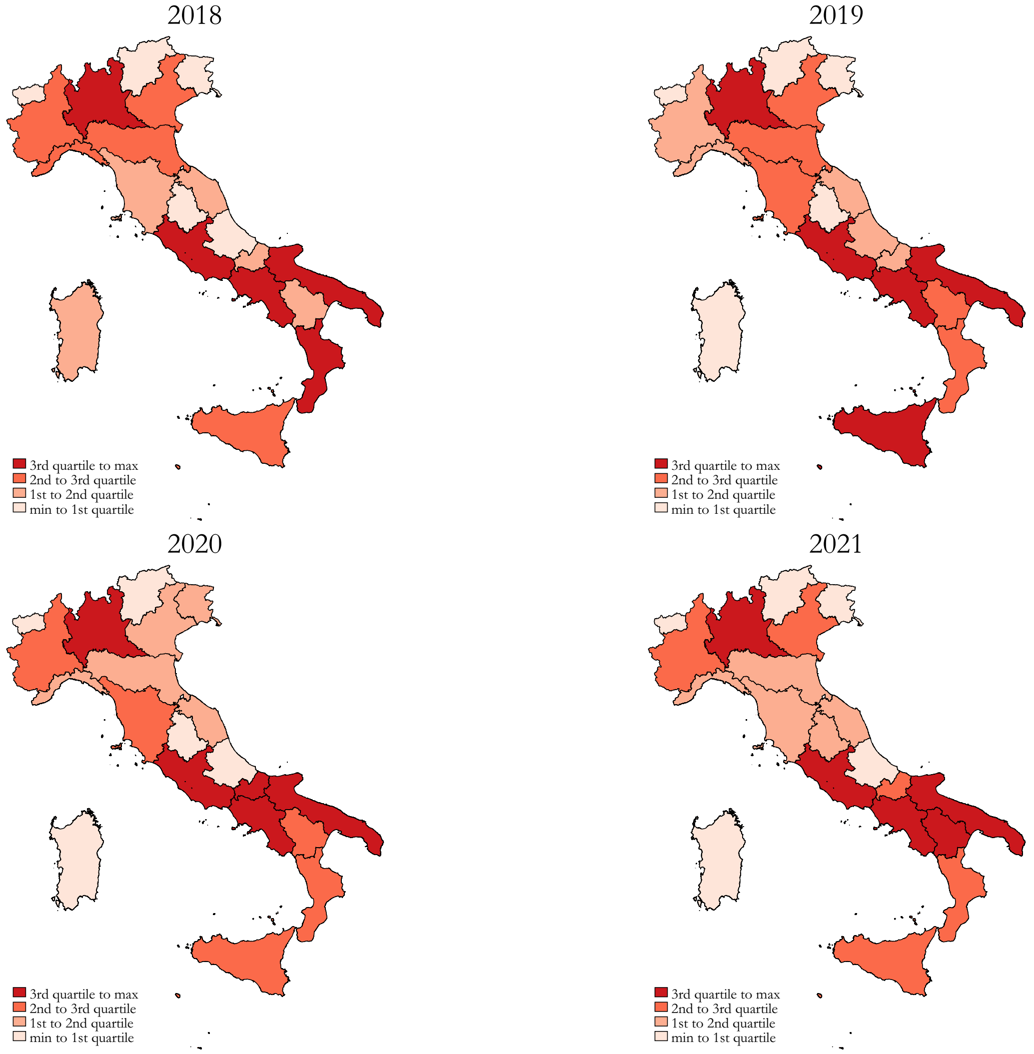
The substantial stability of the annual estimates in relation to GDP suggests that ML is characterised by a 'pro-cyclical' pattern, that is, its magnitude fluctuates in a positive correlation with the economic business cycle. This aligns with theoretical expectations: since ML requires the use of the official economic system to be implemented, the value of financial transactions associated with it tends to follow the broader economic-financial cycle.

[42] The results in Table 3 may be subject to minor adjustments due to potential revisions of past STRs and of SARA data.

By comparing the ML estimates that we derived with the value of suspicious financial transactions observed in reported STRs, one can see that there is a strong asymmetry between the portions of 'omitted' and 'irrelevant' financial values, with the latter aggregate significantly exceeding the former. This is coherent with one of the findings of Gara and Pauselli (2020) which, discussing measurement issues in banks' reporting of suspicious transactions, noted that banks may have an incentive to over-report, as sanctions apply only to omitted reports. Overall, this implies that the estimated ML accounts 'only' between 62% and 70% of total financial transactions reported to UIF by banks.

As detailed above, the proposed methodology involves separate estimations for each of the twenty Italian regions, allowing for a straightforward regional breakdown of the national results. Figure 4 illustrates this regional decomposition for the years 2018 to 2021.[43] The results are fairly consistent over time, with many regions remaining in the same quartile or changing by only one quartile from year to year. The highest share of GDP affected by ML is observed in the most populated regions of Lombardy and Lazio – which are traditional headquarters of private and public entities, respectively – as well as in the Southern regions of Campania, Apulia, Calabria, and Sicily, which are traditionally known for the strongest presence of organized crime.

Figure 4. Regional decomposition of Italian ML (quartiles of ML as % of regional GDP)



[43] The estimation accuracy at the regional level for the most recent year, 2022, may be affected by incomplete investigative interest at the time of data extraction.

6. Concluding remarks

How does ML affect global society? According to Camdessus (1988), money subject to laundering is less productive and, as a consequence, it contributes minimally to a country's economic growth and development. In addition to this, its potential negative impact on the economic system may be decisively bigger. Among adverse effects it is possible to list inexplicable changes in money demand unrelated to economic fundamentals, greater prudential risks to bank soundness,[44] contamination effects on legal financial transactions, and greater volatility of international capital flows and exchange rates due to unanticipated cross-border asset transfers. More in general, but not less importantly, also the social and political implications of ML and predicate offences should be considered, including the overall weakening of the social fabric and collective ethical standards.

Estimating the size of ML remains inherently a challenging task due to the hidden nature of illicit financial flows and limitations in available data and methodologies. Despite intrinsic uncertainty and shortcomings, this paper arguably represents a significant methodological advancement compared to previous evaluations, providing a more reliable estimate of ML for Italy.

It contributes to the existing literature on methods for estimating the size of ML on several dimensions. A novel and comprehensive database, including Suspicious Transaction Reports (STRs) and other financial data from the Italian Financial Intelligence Unit as well as other sources, has been used through the analysis. We also introduce an original methodology consisting of three key phases: rigorous selection of STRs, adoption of a machine learning algorithm to identify reliable information, imputation of unreliable information by a technique of statistical matching in order to ultimately derive a plausible measure of ML. Our findings suggest that ML in Italy accounted on average for approximately 1.8% of GDP between 2018 and 2022, a figure that aligns with official estimates of illegal activities.

The estimates obtained here should be considered as a conservative quantification of the actual magnitude of ML in Italy for several reasons. First, they rely solely on information from suspicious financial transactions reported by banks; including data from other financial intermediaries and non-financial obliged entities would result in a larger estimate. Second, since the imputation process implemented in the third phase is entirely based on reported financial operations observed for bank-province combinations identified as reliable, it excludes cases of ML that are largely unknown or completely unrecognized phenomena (or connected to new types of illicit activities).[45] Lastly, our methodology focuses on the nominal values of the suspicious transactions being detailed in each STR.[46] While each STR provides both the values of such transactions and an option to report the 'total amount of suspicious financial operation' - which typically exceeds the sum of the single transactions but is more prone to measurement errors – considering this total amount could lead to a higher estimate.

One potential enhancement of the current methodology could involve the computation of standard errors for the provided estimates. Incorporating standard errors would not only strengthen the statistical rigor of our approach but also enable us to empirically test a key underlying hypothesis of the methodology. Specifically, it would allow us to assess whether the assumed share of unreliable bank-provinces (currently set exogenously between 10% and 30%) could be adjusted based on empirical evidence. This refinement could lead to a more precise estimate, potentially replacing the current interval range with a single, more accurate, figure.

Understanding the scale and the effects of ML is a crucial objective. A reliable estimate of the phenomenon provides politicians and policymakers with the critical information needed to prioritize

[44] To this respect, he noted how the Core Principle for Effective Banking Supervision (approved by the Basel Committee in September 1997) states that "Banking supervisions must determine that banks have adequate policies, practices and procedures in place, including strict 'know-your-customer' rules, that promote high ethical and professional standards in the financial sector and prevent the bank being used, intentionally or unintentionally, by criminal elements."

[45] Effective examples are illicit activities based on the use of cryptocurrencies or dark web technologies.

[46] As specified in the data section, this serves the dual purpose of allowing a more precise geographical ascription of the suspicious financial transactions and effectiveness matching with the bank-provinces identified in the SARA database.

issues and formulate appropriate policy responses (Ferwerda et al., 2020). It represents a key tool for accurately assessing the phenomenon, identifying its potential drivers, and understanding its consequences on the legal economy. Additionally, it allows the evaluation of trends and patterns, and provides a breakdown of its magnitude at the geographical level, providing risk indicators crucial for establishing an effective risk-based prevention and law enforcement action. Moreover, it aids in evaluating the potential impact of regulatory changes or broader policy shifts on ML activities.

It is also important to differentiate our estimates of ML from other measures of illegal monetary flows. Our estimate is intended to quantify the annual amount of illegal funds that are introduced into the official financial system. This figure may differ significantly from other related metrics, such as the total proceeds of criminal organizations operating or residing in Italy, the laundered aggregate of the total illegal funds, including those earned and reinvested abroad, or the overall wealth accumulated by mafia groups over the years. Although these aspects are also important, they fall outside the scope of our study.

Estimating ML can be useful also from an operational perspective. There are several contexts where the output of this work could be helpful, such as in the periodic Risk Assessment of the national AML system and the Mutual Evaluations conducted by the FATF. By decomposing the size of ML at the level of territorial units, and expressing it as a percentage of regional GDP or any other financial variables, authorities engaged in preventing and combating ML can perform their tasks more effectively and efficiently. This granular approach might enable targeted interventions and resource allocation, enhancing the overall efficacy of AML efforts.

References

Aljassmi, M., Gamal, A.A.M., Abdul Jalil, N., David, J., and Viswanathan, K.K. (2024). Estimating the magnitude of money laundering in the United Arab Emirates (UAE): Evidence from the currency demand approach (CDA). *Journal of Money Laundering Control*, 27(2), 332-347.

Ardizzi, G., Petraglia, C., Piacenza, M., Schneider, F., and Turati, G. (2014). Money laundering as a crime in the financial sector: A new approach to quantitative assessment, with an application to Italy. *Journal of Money, Credit and Banking*, 46(8), 1555-1590.

Ardizzi, G., and Righi, A. (2022), Statistical and Forecasting Use of Electronic Payment Transactions: Collaboration between Bank of Italy and Istat, *Bank of Italy Markets, Infrastructures, Payment Systems Working Paper*, 27.

Argentiero, A., Bagella, M., and Busato, F. (2008). Money laundering in a two-sector model: Using theory for measurement. *European Journal of Law and Economics*, 26, 341-359.

Astarita, C., Capuano, C., and Purificato, F. (2018). The macroeconomic impact of organised crime: A post-Keynesian analysis. *Economic Modelling*, 68, 514-528.

Bagella, M., Busato, F., and Argentiero, A. (2009). Money laundering in a microfounded dynamic model: Simulations for the U.S. and the EU-15 economies. *Review of Law and Economics*, 5(2), 879-902.

Barone, G., and Mocetti, S. (2014). Natural disasters, growth, and institutions: A tale of two earthquakes. *Journal of Urban Economics*, 84, 52-66.

Barone, R., and Masciandaro, D. (2011). Organized crime, money laundering and legal economy: Theory and simulations. *European Journal of Law and Economics*, 32, 115-142.

Breiman, L. (2001). Random forests. Machine Learning, 45, 5-32.

Camdessus, M. (1998). Money laundering: The importance of international countermeasures. Speech to the Financial Action Task Force, at the Plenary Meeting of the Financial Action Task Force on Money Laundering, 10, Paris.

Cariello, P., De Simoni M., and Iezzi, S. (2024). A machine learning approach for the detection of firms linked to organised crime in Italy, based on balance sheet data. *UIF, Quaderni dell'antiriciclaggio, Collana Analisi e Studi*, 22.

Chaudhuri, P., and Loh, W. (2002). Nonparametric estimation of conditional quantiles using quantile regression trees. *Bernoulli*, 8, 561-576.

Cusano, F., Marinelli, G., and Piermattei, S. (2022). Learning from revisions: An algorithm to detect errors in banks' balance sheet statistical reporting. *Quality & Quantity*, 56, 4025-4059.

De Carolis, F., Giorgiantonio, C., and Giovanniello, V. (2014). L'affidamento dei lavori pubblici in Italia: un'analisi dei meccanismi di selezione del contraente privato. *Banca d'Italia, Questioni di Economia e Finanza*, 83. (Available only in Italian).

Ferwerda, J., van Saase, A., Unger, B., and Getzner, M. (2020). Estimating money laundering flows with a gravity model-based simulation. *Scientific Reports*, 10(18552).

Gara, M., and Pauselli, C. (2020). Looking at 'Crying Wolf' from a different perspective: An attempt at detecting banks under- and over-reporting of suspicious transactions. *Italian Economic Journal*, 6(2), 299-324.

Koenker, R. (2005). Quantile Regression. Cambridge University Press.

Koenker, R., Ng, P., and Portnoy, S. (1994). Quantile smoothing splines. *Biometrika*, 81, 673-680.

Lenza, M., Moutachaker, I., and Paredes, J. (2023). Density forecasts of inflation: A quantile regression forest approach. *ECB Working Paper,* No. 2830.

Levi, M., Reuter, P., and Halliday, T. (2018). Can the AML system be evaluated without better data? *Crime, Law and Social Change*, 69, 307-328.

Ling, W., Cheng, B., Wei, Y., Willey, J.Z., and Cheung, Y.K. (2022). Statistical Inference in Quantile Regression for Zero-Inflated Outcomes. Stat Sin. 32(3), 1411-1433.

Little, R. J. A. (1988). Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics*, 6(3), 287-296.

Little, R. J. A. (2013). In praise of simplicity not mathematistry! Ten simple powerful ideas for the statistical scientist. *Journal of the American Statistical Association*, 108(502), 359-369.

Loayza, N., Villa, E., and Misas, M. (2019). Illicit activity and money laundering from an economic growth perspective: A model and an application to Colombia. *Journal of Economic Behavior & Organization*, 159, 442-487.

Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, 7, 983-999.

Mocetti, S. and Rizzica, L. (2023). Organized crime in Italy: An economic analysis. *Italian Economic Journal*, 1-22.

Moiseienko, A., and Keatinge, T. (2019). The scale of money laundering in the UK: Too big to measure? *RUSI Briefing Paper.*

Novaro, R., Piacenza, M., and Turati, G. (2022). Does money laundering inflate residential house prices? Evidence from the Italian provincial markets. *Kyklos*, 75(4), 672-691.

Pinotti, P. (2015). The economic costs of organised crime: Evidence from Southern Italy. *The Economic Journal*, 125(586), F203-F232.

Reuter, P. (2013). Are the estimates of the volume of money laundering either feasible or useful? In B. Unger & D. van der Linde (Eds.), *Research Handbook on Money Laundering* (pp. 159-171). Edward Elgar Publishing.

Rubin D. B. (1986), Statistical matching and file concatenation with adjusted weights and multiple imputations, *Journal of Business & Economic Statistics*, 4, 87-94.

Schneider, F. (2007). Shadow economies and corruption all over the world: New estimates for 145 countries. *Economics - The Open-Access, Open-Assessment E-Journal*, Kiel Institute for the World Economy (IfW Kiel), vol. 1, 1-66.

Schneider, F. (2010). Turnover of organized crime and money laundering: Some preliminary empirical findings. *Public Choice*, 144, 473-486.

Schneider, F. (2013). The financial flows of transnational crime and tax fraud in OECD countries. *Public Finance Review*, 41(5), 677-707.

Schneider, F., and Windischbauer, U. (2008). Money laundering: Some facts. *European Journal of Law and Economics*, 26(3), 387-404.

Tanzi, V. (1980). The underground economy in the United States: Estimates and implications. *Banca Nazionale del Lavoro Quarterly Review*, 135, 427-453.

Tanzi, V. (1983). The underground economy in the United States: Annual estimates 1930-1980. *IMF Staff Papers*, 30, 283-305.

Tanzi, V. (1996). Money laundering and the international financial system. *IMF Working Paper*, 96/55.

Transcrime (2015). Gli investimenti delle mafie, Progetto PON sicurezza, 2007–2013. Transcrime e Università Cattolica del Sacro Cuore, Milano.

UIF (2017), Annual Report, available in English at: https://uif.bancaditalia.it/pubblicazioni/rapporto-annuale/2017/Annual_Report_2016.pdf?language_id=1

UIF (2022), Quaderni dell'antiriciclaggio dell'Unità di Informazione Finanziaria, First semester 2022 (Available only in Italian).

UIF (2023a), Annual Report, available in English at: https://uif.bancaditalia.it/pubblicazioni/rapporto-annuale/2023/annual_report_2022.pdf?language_id=1

UIF (2023b), Quaderni dell'antiriciclaggio dell'Unità di Informazione Finanziaria – Dati statistici," Second semester 2022 (Available only in Italian).

Unger, B. (2013). Can money laundering decrease? *Public Finance Review*, 41(5), 658-676.

UNODC (2011). Estimating illicit financial flows resulting from drug trafficking and other transnational organized crime. Vienna: United Nations.

Van Buuren, S. (2018). Flexible Imputation of Missing Data. Chapman & Hall.

van Duyne, P.C. (1994). Money-laundering: Estimates in fog. *Journal of Financial Crime*, 2(1), 58-74.

Walker, J. (1995). Estimates of the extent of money laundering in and through Australia. Paper prepared for Australian Transaction Reports and Analysis Centre.

Walker, J. (1999). How big is global money laundering? *Journal of Money Laundering Control*, 3(1), 25-37.

Walker, J. (2007). The extent of money laundering in and through Australia in 2004. Criminology Research Council.

Walker, J. and Unger, B. (2009). Measuring global money laundering: The "Walker gravity model". *Review of Law and Economics*, 5(2), 821-853.

Walker, J. and Unger, B. (2013), Measuring global money laundering: The 'Walker gravity model. In B. Unger & D. van der Linde (Eds.), *Research handbook on money laundering*, 159-171. Cheltenham: Edward Elgar.

Yuan, S. (2015), Random gradient boosting for predicting conditional quantiles, *Journal of Statistical Computation and Simulation*, 85(18), 3716-3726.

Zizza, R. (2002), Metodologie di stima dell'economia sommersa: un'applicazione al caso italiano, Bank of Italy, *Temi di Discussione*, 463.