

RESEARCH

Open Access



Modeling the structure and dynamics of transnational illicit networks: an application to cigarette trafficking

Cecilia Meneghini*, Alberto Aziani and Marco Dugato

* Correspondence: cecilia.meneghini@unicatt.it
Università Cattolica del Sacro Cuore and Transcrime, Milan, Italy

Abstract

Many illicit markets are transnational in nature: illicit products are consumed in a country different from the one in which they were produced. Therefore, reconstructing the trafficking network and estimating the size of cross-border illicit flows are crucial steps to gain better understanding of these crimes and to enforce actions aimed at countering them. In this respect, the present study outlines a methodology with which to map and size illicit flows and applies it in estimation of transnational cigarette trafficking flows. The proposed methodology traces each step in the paths followed by illicit cigarettes flowing from their origin to the final consumption country and then estimates the quantity of cigarettes moving between each pair of countries. It exploits data on consumption of illicit cigarettes in 57 countries located in Europe, the Middle East, Central Asia, and North Africa, together with data on seizure cases and geographical information for 158 countries worldwide. These data are combined by a function that assigns a likelihood value of illicit cigarettes being transported between any pair of countries. An algorithm is then implemented in order to identify the most likely paths from the origin to the destination country. By merging results for all the different combinations of origins and destinations, this study estimates the size of all cross-border illicit flows and reconstructs a dynamic transnational cigarette trafficking network for the period 2008–2017. The results highlight the multifaceted role of countries in the cigarette trafficking network, the emergence of identifiable cigarette trafficking routes, and the evolution over time of the structure of this transnational illicit network. Finally, the paper discusses how the methodology developed could be adapted to the study of other transnational crimes.

Keywords: Cigarette smuggling, Illicit market, Illicit flows, Network analysis, Network dynamics

Introduction

Illicit cigarette trafficking is a global and transnational crime causing substantial losses in public revenues, undermining the effectiveness of smoking-reduction policies, and providing financial resources to organized crime and terrorist groups (Joossens et al. 2000; Kulick et al. 2016). It is “perhaps the most widespread [...] sector in the shadow



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

economy” (OECD 2016 p. 16). The threatening size of this illicit market calls for deeper understanding of its actual magnitude and characteristics. However, the paucity of available data on all dimensions of cigarette trafficking is an impediment to research on this topic. There are three main aspects that complicate the measurement of cigarette trafficking activities. The first consists in the illegal nature of this market: criminals hide their actions to protect their loads and to avoid being identified and prosecuted (Joossens and Raw 1998; Gallagher et al. 2019). Similarly, smokers purchasing illicit cigarettes may be unwilling to reveal information about their deviant behavior during interviews and surveys. Second, the dual nature of cigarette consumption (i.e., the coexistence of a legal and illegal market) causes the illegality of a cigarette to depend on its supply chain, which complicates the identification of illegal products (Reuter and Majmundar 2015). Third, the dual market dimension induces public opinion to consider cigarette trafficking as a less significant crime compared to other types of trafficking (e.g. of drugs or firearms) (von Lampe 2002). This creates both scant awareness among consumers of their illicit conduct, and the existence of a relatively low level of enforcement, which contributes to the lack of data on law enforcement operations related to this crime.

Researchers attempting to resolve these issues mostly focus on estimating and analyzing the prevalence of illicit cigarette consumption in certain countries of interest without focusing on the transnational dimension of the market. To the best of our knowledge, there is no work that attempts to estimate the overall network of illicit cigarette flows targeting a comprehensive set of destination countries. Yet, including knowledge on the transnational flows bringing illicit cigarettes to their final destination markets is crucial to effectively quantify a crime which is primarily transnational. Focusing on the overall network structure is essential in order to gain a thorough understanding of the criminal behavior driving these illicit traffics.

Social network analysis approaches are increasingly used in the study of crime. On the one hand, this is due to the recognition of the role that connections among individuals or places have in the understanding of criminal behaviors; on the other hand, it is boosted by the increasing availability of relational data due to new technologies and new areas of analysis (e.g., cybercrime). For example, social network analysis has been used to focus on individuals’ connections and communications in order to map and disrupt criminal networks by identifying the most influential subjects (Calderoni 2014; Ferrara et al. 2014; Bright 2015; Alzaabi et al. 2015; Taha and Yoo 2016). Similarly, it has been applied to the countering of frauds and suspicious activities starting from the analysis of economic transactions or business ownership relations among individuals or legal entities (Fronzetti Colladon and Remondi 2017). However, most of these applications focus on micro- or individual-level networks, while only a few studies analyze criminal connections at the macro-level (Giommoni et al. 2017).

This paper follows this latter strand of research. It contributes to understanding of the illegal trafficking of cigarettes by modeling transnational cigarette trafficking as a dynamic, directed and weighted network in which nodes represent countries and links connecting the nodes are the volumes of illicit cigarettes trafficked between them. We develop a methodology that first combines evidence on seizure cases and illicit cigarette consumption to estimate how likely it is for illicit cigarettes of a specific origin to flow between each pair of countries. Next, we rely on a Yen’s algorithm (Yen 1971) to

identify the most likely routes taken by illegal cigarettes flowing from their countries of production to the markets in which they are consumed. We do so by exploiting information emerging from the previous step as well as data on geographical distance between countries. We then redistribute the volume of illicit cigarettes consumed over the identified paths in proportion to the estimated path-specific likelihood. By aggregating results for all the different combinations of origins and destinations, we are finally able to reconstruct the dynamic cigarette trafficking network. The methodology we propose can potentially be applied in the estimation of trafficking routes of other types of illicit markets, which would make it possible to overcome the central issue in criminal networks research of reconstructing the network of illegal activities on which data are not directly available.

The paper is structured as follows. The next section reviews previous studies dealing with the estimate of trafficking routes of different illicit goods, with a specific focus on illicit cigarettes, and it then introduces our research questions. The third section discusses the empirical approach adopted to estimate the network of transnational illicit cigarette flows, presenting the data used in the methodology and the different steps for the network reconstruction. The fourth section outlines the main results: transnational cigarette trafficking is truly global, and it rapidly changes. The final section discusses limits and potential of the proposed estimation approach before concluding.

Background

In the extant literature, the cigarette trafficking debate mainly focuses on measuring the illicit cigarette markets at the country level rather than on the estimation of transnational illicit flows. While there are some studies that in fact deal with the investigation of trafficking connections between countries (e.g., Joossens and Raw 1998; Mackay and Eriksen 2002; Yürekli and Sayginsoy 2010; Allen 2011), all these analyses focus on a limited set of trafficking relationships and none of them attempts to provide a picture of the overall trafficking network. Furthermore, most of the previous studies adopt a national or regional—especially European Union—perspective, overlooking the existing connections among different regions. In the same vein, etiological studies mostly investigate the determinants of the size of illicit cigarette markets or the reasons why certain countries are primary producers or consumers of illicit cigarettes, rather than the drivers of cigarette trafficking routes. In network terms, this means that existing studies provide estimates at the node level or consider only a few dyadic relationships; but no comprehensive estimates on the relational aspects in the network are produced, even though the transnational nature of cigarette trafficking is widely acknowledged (see, for example, Interpol 2014). Indeed, in the overwhelming majority of cases, illicit cigarettes are smoked in a country different from the one in which they were produced.

The transnational dimension of cigarette trafficking is not an anomaly in the illicit markets domain. Most illicit markets are transnational in nature, meaning that they are operated and/or planned in more than one country, or they are operated in one country but have substantial effects in another, or they are operated in one country but involve an organized criminal group in another one (Interpol 2014). Previous works have sought to estimate the routes used by traffickers in different types of illicit markets, including counterfeits (e.g., OECD, EUIPO 2017), human smuggling (e.g., Tinti 2017), trafficking in radiological and nuclear material (e.g., Sin and Boyd 2016), and drug

trafficking, which includes the largest body of studies (e.g., Boivin 2014; Chandra and Joba 2015; Giommoni et al. 2017; Chandra et al. 2017). However, none of these studies has proposed a systematic method to estimate transnational trafficking routes that could also be applied to illicit cigarette flows, since most of these works exploit specific aspects of the good being trafficked in the estimation methodology. In the case of drug trafficking, for example, studies mapping trafficking routes mainly rely on data on drug seizures, drug consumption and prices, interviews, and geographical dimensions. Compared to trafficking in illicit cigarettes, there is a clearer distinction between countries involved in drug trafficking as points of origin or destination markets for the drugs. This makes it possible to infer a priori the direction of trafficking flows on the basis of drug prices, thus deriving a network that does not allow for the existence of bi-directional flows. The assumption that each pair of countries is not involved in bi-directional flows cannot be applied to cigarette trafficking, which comprises the consumption of different types of illicit products with different across-country price differentials, causing the distinction between origin and destination countries to be dependent on the type of trafficked cigarette. Moreover, some studies focusing on the estimation of trafficking routes exploit known determinants of the trafficking of that specific good to estimate the trafficking network. However, this approach requires reliance on specific assumptions concerning the factors that affect the volume of illicit goods being trafficked to estimate the trafficking network.

As already noted, studies that attempt to describe cigarette smuggling routes provide only some partial or local estimates of these routes, primarily due to data limitations. Joossens and Raw (1998) reconstruct three cigarette smuggling routes in Europe on the basis of newspaper articles and official EU documentation. Mackay and Eriksen (2002) provide a map of some worldwide cigarette trafficking routes, but their estimates are based on data collected prior to 2000 and include only few known routes. Yürekli and Sayginsoy (2010) estimate worldwide smuggling routes by considering some variables identifying smuggling incentives, such as price disparities and corruption levels. They also rely on data on legal exports and imports to identify the major hub/transit countries in each region. Then, they run a simulation analysis to evaluate how changing some variables affects the estimates of smuggling routes. However, their analysis bases the estimation of trafficking routes on some variables that are hypothesized to be drivers of cigarette smuggling, rather than on actual available evidence of cigarettes following a specific route. Finally, Allen (2011) presents a map of the main smuggling routes for brands consumed in the Southern Africa region. Overall, previous studies have only provided a partial picture of the cigarette trafficking phenomenon and most of them are fairly outdated, especially considering the economic, legislative, and political factors that have affected the cigarette market in recent years.

In order to fill the gap in the research on transnational cigarette trafficking, we present a methodology to estimate the entire cigarette trafficking network describing the illicit flows of cigarettes consumed in 57 major destination countries. Besides going beyond the regional perspective, estimates are also longitudinal because they cover a 10-year time span, thus making it possible to explore the evolution of cigarette trafficking over time. Compared to the majority of studies attempting to estimate trafficking routes, we develop a methodology that does not exploit data on the structural determinants of cigarette trafficking (e.g., number of consumers, price incentives, corruption

level). In this way, determinants could be exploited as etiological explanations for the structure of the network, thus making it possible to assess different theories of crime. The proposed methodology could also be applied in the estimation of the network of other types of illicit flows because it does not rely on any cigarette-specific assumption. In addition, assuming the availability of data on illicit cigarette consumption for all the countries of the world, the methodology could be applied to extend the analysis to estimation of the global network of cigarette trafficking.

Given these premises, we address the following research questions:

1. What is the structure of the cigarette trafficking network related to illicit cigarette flows destined to be consumed in 57 countries located in Europe, the Middle East, Central Asia, and North Africa?
2. Is there an evolution over time in the structure of this cigarette trafficking network?

The next section outlines the data and the different methodological steps employed to answer these questions.

Data and methodology

To derive the dynamic, directed and weighted network of transnational cigarette trafficking, this study relies on a three-step methodology that exploits data on a comprehensive sample of seizure cases, illicit cigarette consumption data, and geographic information. The first step uses an econometric model to combine data on illicit cigarette consumption and seizure cases and to compute the weights of '*origin-specific likelihood networks*', i.e. networks in which links represent the likelihood of movements of illicit cigarettes from a given origin.¹ The second step relies on Yen's algorithm (Yen 1971) to search for the paths with highest likelihood from each origin country to each destination country. It exploits information emerging from the first step together with data on geographical distance between countries. As cigarette traffickers utilize multiple pathways to reach a final market, Yen's (1971) approach emerges as an efficient method for introducing, and accounting for, that uncertainty. Then, the volume of illicit cigarettes consumed is redistributed over the identified paths in proportion to the estimated paths-specific likelihood. The third and last step assembles all the derived trafficking networks by summing up the volumes of illicit cigarettes of different origin flowing through the same link, thus reconstructing the aggregate transnational cigarette trafficking network for each year in the 2008–2017 period. The next section describes the employed data; the following one details the three-step methodology.

Data

The proposed methodology relies on different variables that can be ascribed to three categories: illicit consumption, seizure cases, and geographical information (see Table 1).

¹It should be specified that the term "likelihood" does not make any direct reference to maximum likelihood estimation methodologies. It instead refers to the concept of an action, state or event being likely.

Table 1 Description of the data used to estimate the transnational cigarette trafficking network, 2008–2017

Variable	Source	Mean	St. Dev.	Min	Max	N
Illicit cigarette consumption	Authors' elaborations of EPS and legal registered production data	67.8 mn sticks	348.6 mn	1530 Spanish illicit sticks smoked in Bulgaria in 2015	8.8 bn Polish illicit sticks smoked in Germany in 2013	9059
'Unknown-origin' seizures	Authors' data collection from open sources	9.6 mn sticks	58.9 mn	60 sticks seized in 2010 while flowing from North Macedonia to Bulgaria	1.6 bn sticks seized in 2017 while flowing from Tunisia to Libya	930
'Known-origin' seizures	Authors' data collection from open sources	7.1 mn sticks	52.3 mn	120 Greek sticks seized in 2017 while flowing from Greece to North Macedonia	1.2 bn Chinese sticks seized in 2011 while flowing from the Netherlands to Germany	584
Geographical distance	Mayer and Zignago 2011	8430.2 km	4663.6 km	72.7 km between Israel and Palestine	19,781 km between Taiwan and Paraguay	47,742

Note: for illicit consumption and seizure data, descriptive statistics exclude all the 0 values, which were attributed to observations with no estimated illicit consumption or no evidence of seizure

Illicit consumption

The consumption of illicit cigarettes in 57 countries located in Europe, North Africa, the Middle East, and Central Asia has been estimated for each year in the time span 2008–2017 (see the [Supplementary material](#) for the full list of these 57 countries). The estimation of illicit cigarette consumption relies on a methodology that combines information on legal registered production²—providing the overall number of cigarettes sold—and data from Empty Pack Surveys (EPSs)³—providing evidence of the consumption of cigarettes in each country—previously developed and refined by KPMG (2018), and Aziani et al. (2020). EPSs are an unobtrusive method with which to gather data on illicit cigarette consumption since they do not require the direct cooperation of smokers, thus avoiding the risk of being affected by under-reporting and social desirability biases (Webb et al. 1966; Merriman 2010; Reuter and Majmundar 2015). EPS data yield a picture of illicit cigarette consumption that is both geographically broad and easily comparable across different countries. By merging information on volumes and origin of the consumed cigarettes in each country, the methodology makes it possible to obtain country-specific time series data of the volume of illicit cigarettes smoked by country of origin of the cigarette.

Seizure cases

Seizure data cover the years 2008–2017 and were obtained through a comprehensive collection of a sample of seizure cases retrieved from online press articles and press

²Data on the legal registered production in the 57 countries of interest were retrieved from the European Commission (2018) for the EU 28 Member States, and provided by tobacco manufacturers for other countries.

³EPSs consist in collecting discarded cigarette packs from public trash receptacles of from the streets to acquire physical evidence of consumption patterns within a country. Once collected, researchers record the features of the packs (e.g., the brand, the presence and nature of tax stamps), thus identifying the market for which the pack was originally intended and whether it is a genuine brand or a counterfeit (Calderoni 2014; Aziani et al. 2017). EPS and legal registered production data have been transmitted to the authors in the framework of the NEXUS project (2019), which has received funding from PMI Impact, a global grant initiative by Philip Morris International to support projects dedicated to fighting illegal trade and related crimes. PMI Impact did not fund this paper, had no role in its writing, and did not exercise any editorial control.

releases of national law enforcement and custom agencies, where the inclusion criterion was the reference to one of the 57 countries for which consumption data are available. The identification of relevant articles relied on the application of the following query:

(seizure OR seized) AND (tobacco OR cigarette) AND (smuggling OR trafficking OR contraband OR counterfeit OR "smuggled cigarette" OR "smuggled tobacco" OR illicit trade)

The query was translated into all the official languages of each EU country and into Arabic, Belarussian, Russian, Serbian, Turkish, and Ukrainian. The query was then run in two news repositories: LexisNexis⁴ and ProQuest.⁵ We then removed duplicate articles and distinct articles which were evidently referring to the same seizure case. Finally, we scanned the websites of different national and international law-enforcement agencies to look for the availability of press releases and/or annual reports that might contain information on seizure cases not reported by the news, and we added these additional cases to the database.

The data collection was implemented with the aim of obtaining a large sample of all seizures carried out by law enforcement and custom authorities in the area of interest. For each news item reporting evidence of a transnational seizure case, we coded details on the origin country of the illicit cigarettes seized (when available), the countries from which cigarettes were outflowing and to which they were inflowing, the year of the seizure, the type of product (i.e. boxed cigarettes or other tobacco products),⁶ and the volume seized. Seizure cases not referring to transnational trafficking were excluded from the analysis. The final database comprised 2496 relevant seizure cases referring to 491 transnational links connecting 96 countries worldwide.

For the purpose of the network reconstruction, seizure data can be divided into two categories: 'unknown-origin' seizures and 'known-origin' seizures. 'Unknown-origin' seizures were retrieved from news items that reported only the information on the countries from and to which the cigarettes were flowing (the 'outflowing country' and the 'inflowing country', respectively). 'Known-origin' seizures, on the other hand, refer to seizure cases for which the news item reported both information on the outflowing and inflowing countries, in addition to details on the country of origin of the illicit cigarettes confiscated (see Table 2).

Seizure cases are the data most commonly employed to estimate trafficking activities as they provide a direct and transparent measure of different dimensions (e.g., quantity, type of product) of trafficking (Boivin 2011). However, as it will be thoroughly explained in the *Discussion and conclusions* section, the use of seizure data as an evidence of trafficking flows entails some potential biases. In particular, detected seizures provide a picture of trafficking flows that might be biased by the action of law enforcement agencies, and in turn collected news items on seizure cases might also be biased due to uneven media coverage of seizure cases and to possible imbalances in the collection of available articles. Nonetheless, given the general lack of information on trafficking

⁴<https://www.lexisnexis.com/en-us/gateway.page>.

⁵<http://www.proquest.com>.

⁶To improve the size of the collected sample of seizure cases, seizure of tobacco products other than boxed cigarettes were also included and their volume (expressed as weight) was converted into cigarette sticks.

Table 2 Example of seizure data structure

Seizure type	Origin country	From (outflowing)	To (inflowing)	Year	Product type	Volume (sticks)
Unknown-origin	Not available	ESP	FRA	2011	Boxed cigarettes	134,000
Known-origin	BLR	POL	DEU	2014	Boxed cigarettes	23,000

activities, seizure data tend to be considered as the less unsatisfactory source of information available (Boivin 2011; UNODC 2015).

Geographical information

Geographical information consists of a variable reporting the geographical distance between two countries. This variable is obtained by computing the weighted distance between the largest cities of any two countries, employing as weights the share of the city in the overall country's population (Mayer and Zignago 2011). This is a way to approximate the distance that illicit cigarettes need to cover to move between different countries, given that information on the specific locations from and to which illicit cigarettes flow is rarely available. Authors investigating the transnational trafficking networks of different illegal drugs have relied on this proxy as well (e.g., Berlusconi et al. 2017; Giommoni et al. 2017; Aziani et al. 2019a). A possible drawback of this measure is that it does not account for the presence of transportation networks that may facilitate the transnational exchange of goods. While for the methodology that we propose the goal was to include a straightforward measure of geographical convenience, as will be presented in step 2 of the methodology, future extensions might be considered to build a more comprehensive indicator of geographical convenience by including further dimensions of transnational distance.

Network construction

Illicit cigarettes originate in many different countries around the world. For instance, in the Netherlands, data on illicit cigarette consumption show that smokers consume illicit cigarettes produced in as many as 82 different countries. Therefore, the volume of illicit cigarettes trafficked from country i to country j consists of cigarettes originating in O different countries, of which one is country i itself, while i may also act as transit country for the cigarettes flowing from the remaining $O - 1$ countries of origin.⁷ The structure of these trafficking networks varies depending on the origin o of the illicit

⁷Illicit cigarettes produced and consumed in the same country are excluded from the computation since they fall outside the scope of this study, which specifically focuses on transnational trafficking. Illicit cigarettes produced and consumed in the same country nonetheless constitute only a very residual part of the illicit factory-made cigarette market in the countries considered, as empirically observed by several studies (e.g., Transcrime 2016; KPMG 2018). Indeed, due to the *modi operandi* behind cigarette trafficking, which has predominantly a transnational dimension, trafficked cigarettes do not end up being consumed in the same country where they have been produced after leaving it (von Lampe 2011; Reuter and Majmundar 2015). The contraband of genuine factory-made cigarettes involves cigarettes intended for legal consumption either being diverted toward the black market of a different country—especially while ‘in transit’—(i.e., large-scale smuggling) or smuggled by networks of loosely affiliated criminals from lower-tax to higher-tax regimes after being fully taxed (i.e., bootlegging) (Joossens et al. 2000; Reuter and Majmundar 2015). The production of counterfeits has been observed in few of the 57 considered countries (Aziani and Dugato 2019), China being by far the largest producer of counterfeits in the world (Swami et al. 2009; Shen et al. 2010). Therefore, also the trafficking of these cigarettes has predominantly a transnational dimension, at least in the countries considered. Finally, illicit whites are factory-made cigarettes that are legally produced by registered enterprises but intended for the illicit market usually outside the jurisdiction where the cigarettes are manufactured (Reuter and Majmundar 2015; Ross et al. 2015).

cigarettes; for example, it might be that traffickers move Korean cigarettes from Libya to Italy, but do not exploit the same path to smuggle Slovenian cigarettes. Accordingly, we model transnational cigarette trafficking as a combination of O ‘origin-specific’ trafficking networks, one for each possible country of origin of the illicit cigarettes. In each of these networks nodes represent countries and the links connecting the nodes are the volumes of illicit cigarettes trafficked between them. Each ‘origin-specific’ trafficking network can be represented by the following matrix:

$$W^o = \begin{bmatrix} w_{11}^o & \dots & w_{1n}^o \\ \dots & \dots & \dots \\ w_{n1}^o & \dots & w_{nn}^o \end{bmatrix} \tag{1}$$

where w_{ij}^o is the volume of illicit cigarettes of origin o flowing from i to j , and $i = \{1, \dots, n\}$, $j = \{1, \dots, n\}$. Here, i is defined as the ‘outflowing’ country and j is the ‘inflowing’ country. In this matrix, the w_{ii}^o element is always 0 since self-loops are not allowed in consideration of the fact that our focus is transnational trafficking and that the very great majority of illicit cigarettes consumed in each country come from abroad.

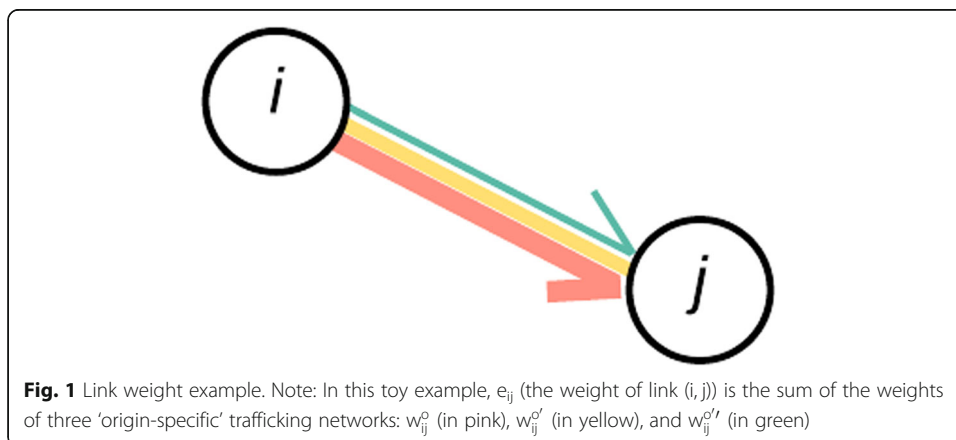
Given that the final aggregate cigarette trafficking network is the sum of all the ‘origin-specific’ trafficking networks, it can be represented by a set $C = (V, E)$, such that $V = \{v_1, v_2, \dots, v_n\}$ are the country nodes, and $E = \{(i, j)\}$ denotes network links (or edges). Each link (i, j) has a weight e_{ij} resulting from the sum of the weights of all the ‘origin-specific’ trafficking networks, i.e.:

$$e_{ij} = \sum_{o=1}^O w_{ij}^o \tag{2}$$

where $o = 1, \dots, O$ are all the possible countries of origin of illicit cigarettes (see Fig. 1). e_{ij} quantifies the aggregate volume of cigarettes trafficked from country i to j . Hence, the aggregate cigarette trafficking network can be represented by the following network matrix:

$$E = \begin{bmatrix} e_{11} & \dots & e_{1n} \\ \dots & \dots & \dots \\ e_{n1} & \dots & e_{nn} \end{bmatrix} \tag{3}$$

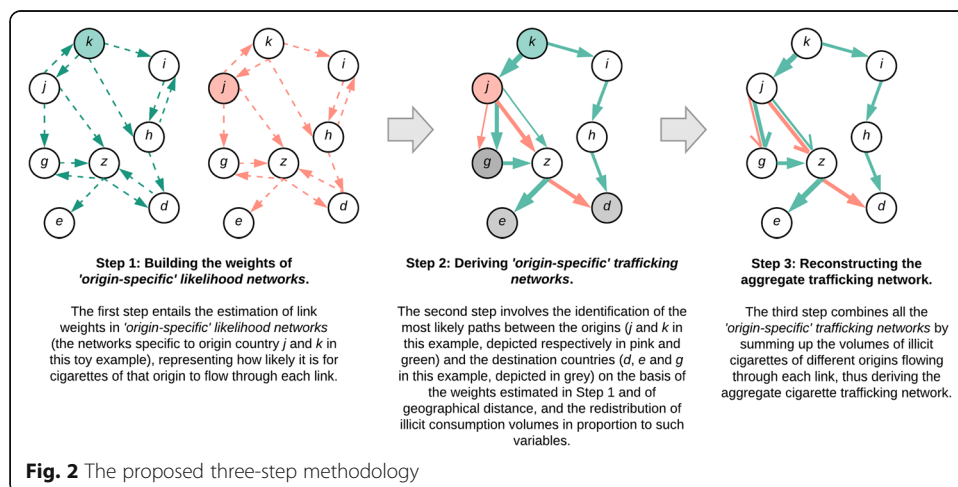
where, as in (1), the e_{ii} element is always 0.



The resulting aggregate trafficking network represents the trafficking relationships exploited by criminals to smuggle cigarettes consumed in 57 countries located in Europe, the Middle East, Central Asia, and North Africa. This geographical scope includes all the major markets of illicit cigarettes consumption in the European area and in neighboring regions. To reach their final destinations, illicit cigarettes originating in 128 different countries travel through countries other than these 57 final markets, thus resulting in a network involving in total 158 countries. The aggregate cigarette trafficking network is both weighted and directed. Accounting for the direction of trafficking relationships is fundamental in order to analyze cigarette trafficking. Indeed, geographical, political, and socioeconomic factors (including but not limited to the characteristics of the legal cigarette market) cause some countries to be key producers of illicit cigarettes—i.e., Belarus—, while some others are mainly or only destination markets for these products—i.e., Norway. The network is also weighted as the estimated flows reflect the volumes of illicit cigarette consumption that vary considerably across countries. Moreover, the estimated network is dynamic, since available data enable the construction of ten different networks, one for each year t in the time span 2008–2017.

The proposed methodology can be divided into three sequential steps: (1) building the weights of ‘origin-specific’ likelihood networks; (2) deriving ‘origin-specific’ trafficking networks; (3) reconstructing the aggregate trafficking network. For better understanding of the process, a summary of the 3-step methodology is presented in Fig. 2.

As stated, the final goal of the methodology is to derive the aggregate cigarette trafficking network, i.e. to derive the values that populate the matrix in (3). The aggregate cigarette trafficking network combines all the ‘origin-specific’ trafficking networks by summing the volumes of illicit cigarettes of different origins flowing through each (i, j) link, as described in (2). This will be detailed in Step 3 of the methodology, “Reconstructing the aggregate trafficking network”. Hence, we first need to derive the matrix in (1) for all possible countries of origin. In order to do this, we rely on Yen’s algorithm (Yen 1971) to search for the most likely paths connecting each country of origin of illicit cigarettes o to their final country of destination d . The combination of EPSs and cigarettes release for consumption data provide information on the volume of illicit cigarettes smoked in each country (the destination country d) by their country of origin o . These volumes are redistributed along the identified paths in proportion to the

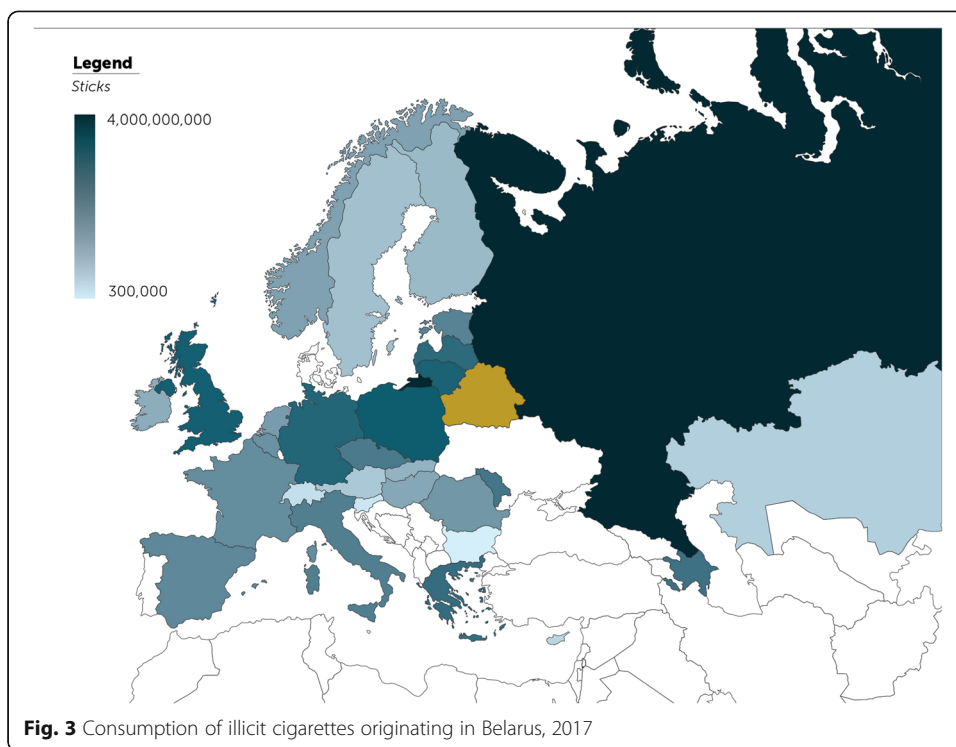


computed likelihood of cigarettes flowing through them. Then, each path is decomposed into its (i, j) steps, and trafficked volumes are summed for each (i, j) link used in multiple paths, thus deriving w_{ij}^o . This procedure is Step 2 of the Methodology, “Deriving ‘origin-specific’ trafficking networks”. To allow Yen’s algorithm (Yen 1971) to search for the most likely paths, we need to build a ‘origin-specific’ likelihood network, where the weight of each link represents how likely it is for illicit cigarettes of a specific origin to travel through it. The estimation of these *origin-specific likelihood networks* constitutes Step 1 of the methodology, which we present in the next subsection.

Building the weights of ‘origin-specific’ likelihood networks

To search for the most likely paths from the country of origin of illicit cigarettes to their final destination country, we first need to determine the relative likelihood of each directed link, i.e. how “plausible” it is for illicit cigarettes to travel from country i to country j . We refer to this likelihood as l_{ij}^o . To estimate this value, we rely on available variables that provide hints that illicit cigarettes are flowing between two countries or being consumed in a specific country: cigarette seizures and consumption of illicit cigarettes. In particular, seizure cases provide direct evidence of the existence of trafficking between two considered countries (Boivin 2014; Giommoni et al. 2017). Similarly, the registered consumption of illicit cigarettes in a specific country testifies that those cigarettes do actually flow through that country. We assume that countries consuming a larger amount of illicit cigarettes of a certain origin are also more likely to be transits for cigarettes of the same origin. Indeed, previous studies have observed strong similarity in the rate of consumption of illicit cigarettes in neighboring countries (Calderoni et al. 2017), and EPS data highlight also the existence of geographical clusters with respect to the origin of illicit cigarettes consumed (e.g., see Fig. 3). This suggests the presence of “spillover effects” in the route chosen by traffickers to ship illicit cigarettes to a certain final destination.

Two types of data on seizures are available—i.e., seizures of cigarettes of ‘known-origin’ and ‘unknown-origin’—depending on the capacity of law enforcement and custom agents to identify the actual country of origin of the cigarettes. ‘Known-origin’ seizure cases provide information on the likelihood of link $(i, j)_{ot}$ being used to transfer cigarettes of origin o in year t : indeed, ‘known-origin’ seizure data give evidence on the out-flowing and inflowing countries but also on the origin of cigarettes being shipped in a specific year. However, ‘known-origin’ seizures do not provide full information on all paths used; they refer only to specific links. Moreover, they constitute only a fraction (approximately 38%) of the total amount of data collected on seizures, and relying on this data alone to guide the mapping of illicit cigarette flows would result in the waste of information provided by ‘unknown-origin’ seizures (see Fig. 4). Hence, we exploit the correlation between ‘known-origin’ seizures and the other two variables—‘unknown-origin’ seizures and consumption—to model the likelihood of illicit cigarettes flowing through links for which we do not have available evidence of trafficking. The idea is to assess how well ‘unknown-origin’ seizures and consumption data explain the volume of ‘known-origin’ cigarettes detected at a particular link in the network, and exploit this information to build the weights of ‘origin-specific’ likelihood networks.



Technically, let us consider link $(i, j)_{ov}$, that is, the directed link between country i and country j in the network of illicit cigarettes with origin o in year t . For the reasons outlined above, the likelihood of this link, l_{ijt}^o , can be modelled as a function of seizure and consumption:

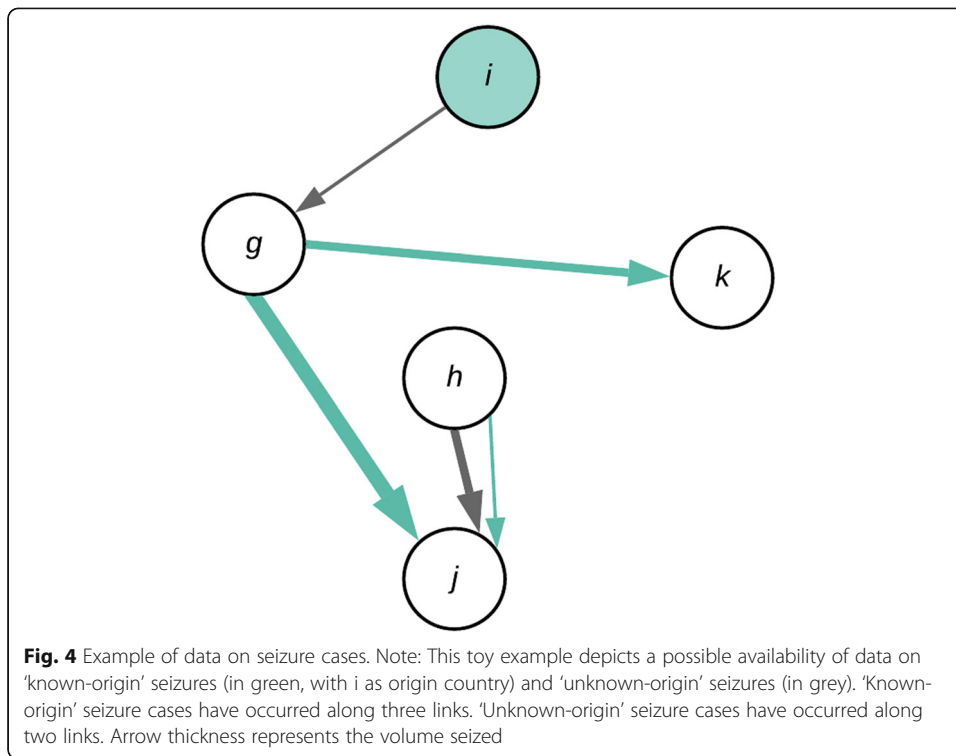
$$l_{ijt}^o = f(seiz_{ijt}, cons_{ojt}) \tag{4}$$

where $seiz_{ijt}$ is the volume of ‘unknown-origin’ illicit cigarettes seized in year t while they were being transferred from country i to country j , and $cons_{ojt}$ is the volume of illicit cigarettes of origin o consumed in country j in year t . The variable $seiz_{ijt}$ provides evidence on the extent of cigarette trafficking from i to j , although no information is available on the origin of these cigarettes, while $cons_{ojt}$ indicates the amount of cigarettes originating in o that reached country j to be smoked there.

We now rely on an econometric approach to find the structure of the function f in (4). We exploit ‘known-origin’ seizures as the dependent variable in a linear regression model where ‘unknown-origin’ seizures and consumption are the independent variables to assess how well these last two items of information explain the volume of ‘known-origin’ illicit cigarettes seized, and thus estimate the function expressed in (4). We then use the information on the explanatory power of ‘unknown-origin’ seizures and consumption to build the likelihood values l_{ijt}^o . In technical terms, we run the following Ordinary Least Squares (OLS) regression model:

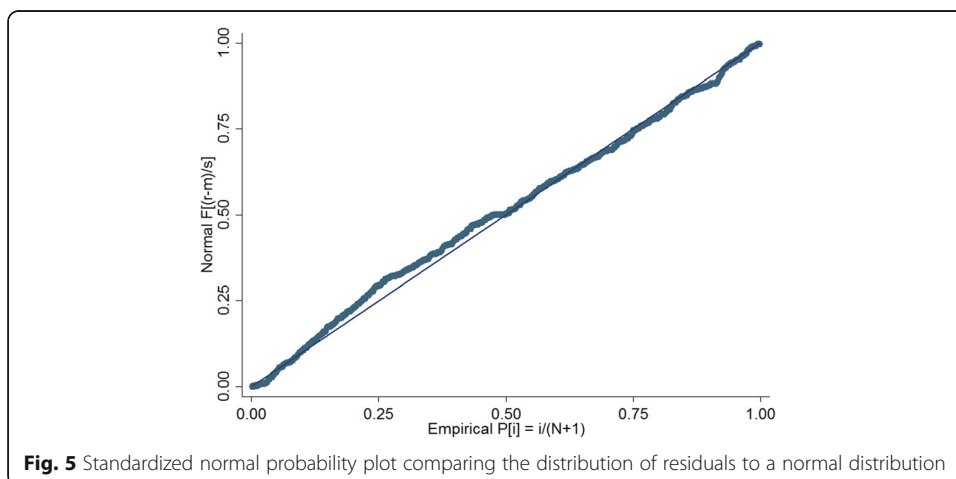
$$\ln (seiz_{ijot}) = \beta_0 + \beta_1 \cdot \ln (seiz_{ijt}) + \beta_2 \cdot \ln (cons_{ojt}) + \alpha_o + \varepsilon_{ijot} \tag{5}$$

where, besides the variables already introduced, $seiz_{ijot}$ is the volume of ‘known-origin’ illicit cigarettes of origin o seized in year t while they were being transferred from



country i to country j , β_0 is the intercept, β_1 is the coefficient representing the relative impact of ‘unknown-origin’ seizures in explaining the volume of ‘known-origin’ seizures, β_2 is the coefficient representing the relative impact of consumption in explaining the volume of ‘known-origin’ seizures, α_o is the unobserved time-invariant origin country fixed effect, and ε_{ijot} is the error term.

Variables are inserted in the model after a natural logarithmic transformation to approximate the skewed distribution of seizure and illicit consumption data better to the normal one. The choice of relying on a OLS regression model ensues from three main reasons: first, the data available satisfy all the assumptions of OLS regression models (e.g., see Fig. 5 for a post-estimation diagnostics showing no major deviations of the distribution of residuals from a normal one). Second, OLS models allow for more



parsimonious though less accurate specifications, which is a preferred choice given the possible imprecisions linked to seizure data. Third, an OLS regression permits the inclusion of country-specific fixed effects in the model. The insertion of country-specific fixed effects makes it possible to control for any biases in the consumption and seizure information collected, while estimating a sufficiently parsimonious model. Biases may be present also at the link level, but the addition of controls at link level would be particularly data-intensive. Moreover, the data collection is unlikely to introduce major biases at link-level, as it is conducted at country-level, and previous studies in the field of drug trafficking have argued that major differences in the effectiveness of law enforcement are unlikely with respect to different links involving the same country (Berlusconi et al. 2017; Giommoni et al. 2017).

The estimated coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$ are employed to construct the link likelihood values (the ones exploited to guide the algorithm searching for plausible paths) of ‘origin-specific’ likelihood networks. More specifically, the results show that the estimated impact of ‘unknown-origin’ seizures in explaining the volume of ‘known-origin’ illicit cigarettes seized is approximately twice the impact of consumption (see Table 3). Both of the estimated coefficients are statistically significant: seizure cases at the 0.1% level, illicit consumption at the 1% level. The estimated relative coefficients of ‘unknown-origin’ seizures and illicit consumption are extrapolated to estimate l_{ijt}^o as $l_{ijt}^o = \frac{0.078}{0.078+0.036} \cdot seiz_{ijt} + \frac{0.036}{0.078+0.036} \cdot cons_{ojt}$. In fact, our concern is not to forecast the volume of ‘known-origin’ seizures related to links for which we do not available data on this type of seizures, but to estimate the likelihood of a given connection being used to traffic cigarettes given the known values of consumption and ‘unknown-origin’ seizures in the two countries.

As a result of this second step, the estimate value of l_{ijt}^o is assigned to the various $(i, j)_{ot}$ links to build the ‘origin-specific’ likelihood networks on which Yen’s algorithm runs, as explained in the next subsection.

Deriving ‘origin-specific’ trafficking networks

The second step of the methodology relies on Yen’s algorithm (Yen 1971) to determine the most likely paths from the country of origin of the illicit cigarettes to their final destination country. It does so by exploiting evidence on consumption and seizures of

Table 3 OLS regression results

	ln(‘known-origin’ seizures)
ln(‘unknown-origin’ seizures)	0.078 (0.014)***
ln(illicit consumption)	0.036 (0.014)**
Intercept	9.996 (0.864)***
Origin country FE	Yes
R ²	0.47
N	584

Notes: Standard errors reported in parentheses. ***, **, and * indicate that the estimated coefficient is statistically significant at the 0.1, 1, and 5% level, respectively

cigarettes and considering also the geographic distance between countries. This step of the methodology resembles the procedure proposed by Sin and Boyd (2016) to discover potential routes of trafficking in radiological and nuclear materials. Sin and Boyd's procedure (Sin and Boyd 2016) uses Dijkstra's algorithm (Dijkstra 1959) to identify paths that minimize the cost for the trafficker or, in other words, maximize the trafficker's chances of success. We choose instead to rely on Yen's algorithm (Yen 1971) so that we can retrieve not only the most efficient path, but the k most efficient—shortest, in Yen's terminology—paths between every pair of nodes. Information provided by seizure cases indicate, in fact, that traffickers are likely to use more than one path to ship illicit cigarettes from the origin to the destination country. Yen's algorithm performs the search in all the '*origin-specific*' *likelihood networks* constructed by exploiting the weights derived in the previous step. Then, to identify paths that converge on the actual final destination d where illicit cigarettes are smoked, the algorithm replicates the search in a second network in which weights are given by the geographical distance between countries. Results from the two types of networks are then merged to obtain a single list of the most likely paths used to traffic illicit cigarettes of a specific origin.

The '*origin-specific*' *likelihood networks* on which the algorithm performs its search are built as follows. First, a directed link is established between each pair of countries if at least one of the following two conditions applies: (1) the two countries share a land border or (2) there is evidence of a seizure case between the two countries in any year included in the analysis (2008–2017).⁸ Second, the networks are weighted by assigning the likelihood value l_{ijt}^o to each established link. The likelihood value l_{ijt}^o has been computed in the previous step relying on information on cigarette consumption and seizures. As consumption values vary on the basis of the 'origin country'-'inflowing country' (o, j) pair and seizure values differ for each 'outflowing country'-'inflowing country' pair (i, j), this weighting procedure results in the creation of multiple '*origin-specific*' *likelihood networks* (on average $n_o = 920$ per year) that differ according to the country of origin of the illicit cigarettes. Finally, Yen's algorithm (Yen 1971) is run over these networks to find the k most likely paths from the origin node to the destination node.⁹

Yen's algorithm relies on Dijkstra's algorithm (Dijkstra 1959) to find the most likely path between two nodes,¹⁰ and then exploits an iterative process to remove paths that have already been explored and find the other $k - 1$ most likely paths. A priori, we do not know how many paths traffickers use to ship illicit cigarettes from the country of production to the destination. We thus run the algorithm for $k = 100$ and look at the stored likelihood values for each path identified. It emerges that some of the retrieved

⁸By "a seizure case between the two countries" we mean that an illegal shipment has been intercepted while it was being moved from country i to country j . In this case, the (i, j) link is established, while the link in the opposite direction is created only in the case of evidence of a seizure case from country j to country i (or if the two countries share a land border).

⁹Technically, Yen's original algorithm is implemented on networks where link weights represent distances and the algorithm thus looks for the shortest paths in the network. Hence, the value of network weights l_{ijt}^o has been inverted for the implementation of this algorithm. In the context of network science and graph theory, the commonly used term 'distance' does not necessarily refer to physical characteristics of a network. Rather, it refers to the more general pairwise relation between two objects. In this study, the intensity of such relation is estimated using data on cigarette consumption, seizures, and geographical distance between countries.

¹⁰Dijkstra's algorithm identifies the shortest path to a destination through an iterative process that explores the candidate neighboring node set and stores the distance information on explored paths (Dijkstra 1959).

paths have a very low estimated likelihood to be used; actually, for some origin-destination combinations Yen's algorithm retrieves less than 100 paths.

To avoid retrieving paths that are very unlikely to be used in comparison to the most plausible one, we decide to discard all paths that are $x = 10$ times less likely compared to the best one. We ran some sensitivity tests to check for the impact of changing x on results. We selected different values of x ($x = 5, 15, 20$) and implemented the following steps of the methodology. Since the remaining part of the methodology involves the re-distribution of consumption volumes over the paths in proportion to their likelihood of being used, the choice of x does not have a substantial impact on the final results, especially for higher values of x . This is because infrequent paths would be assigned to very low consumption volumes in any case. In particular, the distribution of trafficked volumes changes by 12.65% for $x = 5$, and by 0.08% for both $x = 15$ and $x = 20$. The choice of $x = 10$ has the purpose of avoiding an excessive distortion of the outcome of the algorithm while at the same time allowing very implausible paths to be dropped. Capping the number of possible paths on the basis of their relative "cost" for traffickers finds strong theoretical and empirical support in the criminological literature. In particular, the rational theory of crime (Cornish and Clarke 1987) and its implications on organized (e.g., Reuter 1983) and transnational crimes (e.g., Guerette and Bowers 2009) suggest that cigarette traffickers are likely to move along the paths maximizing their profits while minimizing their costs. Accordingly, empirical studies dealing with other transnational traffics also highlight a high concentration of illicit goods flowing along few transnational paths (e.g., Boivin 2014; Chandra and Joba 2015; Berlusconi et al. 2017; Savona and Mancuso 2017; Aziani et al. 2019a).

We then incorporate information on geographical distance to identify paths that converge on the actual final destination d where illicit cigarettes are smoked. Hence, we run Yen's algorithm on networks that have the same structure as the '*origin-specific likelihood networks*' incorporating seizure and consumption data, but where link weights represent the geographical distance between countries. More specifically, each link weight $l'_{ij,d}$ is built as follows:

$$l'_{ij,d} = dist_{ij} + dist_{j,d} \quad (6)$$

where $dist_{ij}$ indicates the distance between the two countries connected by that link, and $dist_{j,d}$ is the distance between the inflowing country j and the final destination of the cigarettes d . This last variable is introduced with the aim of guiding the algorithm in choosing paths that bring cigarettes closer to their final destination. In this respect, the inclusion of geographical variables makes it possible also to add a destination-specific dimension to the procedure. Geographical variables could not have been added to the previous step of the methodology (the determination of weights in '*origin-specific likelihood networks*') since the impact of distance in predicting cigarette trafficking flows cannot be estimated through the use of seizure data. Media are more likely to report news on large seizures, which usually characterize transnational shipments coming from distant locations. Hence, the role of geographical distance in predicting the volume of 'known-origin' seized cigarettes would likely be biased.

More in general, other works focusing on transnational illicit traffics highlight the imperfect nature of data on illicit flows, which cause a necessary trade-off between the

use of incomplete but accurate anecdotal data and making reasonable assumptions that make it possible to overcome the incompleteness of flow data (e.g., Paoli et al. 2009; Chandra and Joba 2015 with respect to illicit drugs). The use of geographical information is based on the assumption that, according to the illegal enterprise theory, illicit trafficking shares some basic principles of geographical convenience with licit transports (Reuter 1983). Scholars, focusing on drug trafficking, indicate that the possibility to bribe custom officers, the presence of ethnic and kinship ties in foreign countries, and the attitude of law enforcement agencies impact on the configuration of transnational trafficking routes (Akyeampong 2005; Desroches 2007; Decker and Townsend-Chapman 2008; Paoli and Reuter 2008). However, other factors being equal, traffickers will choose to follow the fastest route to go from country i to j , in consideration also of the fact that shortest routes tend to be less risky due to the lower number of border crossings and consequently of custom controls as theoretically and empirically argued in previous studies (Reuter 2014; Berlusconi et al. 2017; Aziani et al. 2019a).

Yen's algorithm is run in the *likelihood networks* of geographical distance following the same procedure employed for the networks incorporating seizure and consumption data (i.e., searching for the 100 most likely paths and then discarding all paths that are 10 times less likely than the best one). Then, the two outcomes are merged by summing up the weights referring to the same link. In this way, we are able to correct Yen's algorithm outcome by an element of geographical convenience, and discard or decrease the relevance of paths that are too unlikely in terms of geographical distance.¹¹ This step of the methodology thus provides a list of plausible paths used to ship illicit cigarettes of a given origin to each country where they are consumed, along with a value for each path representing the likelihood that cigarettes of that given origin-destination combination flow through it. We choose to apply the geographical correction by summing the likelihood values arising from geographical distance networks with those derived from the '*origin-specific*' *likelihood networks* as the resulting relevance of the identified paths is well corroborated by similar results emerging from qualitative studies (e.g., Chionis and Chalkia 2016; Johnston et al. 2016; CSD 2018). Nonetheless, as we will argue in the *Discussion and conclusions* section, future works should focus on developing more rigorous data-driven approaches to incorporate geographic information in the estimation of link-specific likelihoods.

Reconstructing the aggregate trafficking network

To reconstruct the entire trafficking network, we redistribute the volume of illicit cigarettes originating in o and consumed in d among all the plausible paths identified in the previous step. Volumes are redistributed proportionally to the estimated likelihood

¹¹As an example, let us consider the trafficking paths used to ship illicit cigarettes originating in Greece and consumed in Germany in 2017. When Yen's algorithm is run on the '*origin-specific*' *trafficking networks* for 2017, the identified likelihood that cigarettes will flow along the path "Greece (GRC) → Turkey (TUR) → Germany (DEU)" is higher than the likelihood that cigarettes will follow the path "Greece (GRC) → Italy (ITA) → Germany (DEU)". This is because seizure and consumption data indicate that the connections "GRC → TUR" and "TUR → DEU" are jointly more likely to be used to ship Greek cigarettes compared to the combination of the connections "GRC → ITA" and "ITA → DEU". However, from a geographic point of view, it is more convenient to use the "Greece (GRC) → Italy (ITA) → Germany (DEU)" path. Indeed, the likelihood that cigarettes will follow this path in the geographical networks is more than twice the likelihood that cigarettes will follow the "Greece (GRC) → Turkey (TUR) → Germany (DEU)" path.

of each path. Then, each path is broken down into all its (i, j) steps, and trafficked volumes are summed up for each (i, j) link to obtain the final weights of ‘*origin-specific*’ trafficking networks. In particular, the same (i, j) connection might be used in different paths departing from the same origin country o . All the estimated volumes flowing through each link of this path are added up to derive w_{ij}^o , i.e. the weight of link (i, j) in the trafficking network of illicit cigarettes originating in o .

Finally, the matrix describing the entire cigarette trafficking network is computed by merging all the ‘*origin-specific*’ trafficking networks. Technically, this merger is achieved by summing all the volumes of cigarettes of different origins flowing through each (i, j) link:

$$E = \begin{bmatrix} \sum_{o=1}^O w_{11}^o & \cdots & \sum_{o=1}^O w_{1n}^o \\ \sum_{o=1}^O w_{n1}^o & \cdots & \sum_{o=1}^O w_{nn}^o \end{bmatrix} \quad (7)$$

where O is the total number of origin countries in the network as identified in the Empty Pack Surveys—i.e., 128 countries.

Results

The aggregate network representing the trafficking of illicit cigarettes consumed in 57 countries located in Europe, the Middle East, Central Asia, and North Africa involves 158 countries that are connected through almost 900 different links (see Table 4, Fig. 6, and Fig. 7 depicting the 2017 network).

The results highlight the global and complex nature of cigarette trafficking. The cigarette trafficking network has a far-reaching structure: some of the illicit cigarettes consumed in the 57 markets of interest originate in geographically very distant countries in East Asia or South America. In many cases, the results show that the preferred routes are not the shortest ones in terms of geographical distance, and illicit cigarettes cross different borders or travel long distances by sea before reaching their final consumption country, even though less intricate routes are available in principle. This finding is in line with those of previous studies that highlight the distinctive *modi operandi* characterizing cigarette trafficking. Differently from other illicit products, cigarettes are per se a legal product, and the diversion to the illegal market can take place at different phases in the supply chain. For example, some cigarettes might be legally produced and exported, and then they might be illegally imported in another country where proper taxes are not paid, or they may end up in a free-trade zone where they are falsely labeled as a different type of good and smuggled to the destination country. Traffickers exploit these grey areas of the legal market to ship their loads, so that cigarettes “disappear” while being transferred to their final destination and authorities lose track of them (von Lampe 2011; Reuter and Majmundar 2015).

Another result highlighting the complexity of this illicit network is that trafficking relationships between pairs of countries often consist of cross-border illicit flows moving in both directions. This is in contrast with findings from studies focusing on the trafficking of other types of illicit goods, as for example drug trafficking or human trafficking flows, where there is a clearer distinction between source and destination countries and thus the presence of mostly unidirectional flows (Boivin 2014; Campana 2018).

Table 4 Network statistics, 2008–2017

Network measure	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Number of involved nodes	155	152	157	158	157	156	157	158	157	157
Edge count	859	854	874	886	891	884	881	892	889	868
In-degree centralization	0.026	0.017	0.019	0.016	0.014	0.033	0.017	0.023	0.023	0.018
Out-degree centralization	0.018	0.014	0.010	0.010	0.008	0.021	0.010	0.018	0.017	0.017
Weighted density	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.002	0.002	0.001
Network assortativity	0.033	0.081	0.186	0.216	0.175	0.115	0.057	0.101	0.103	0.107

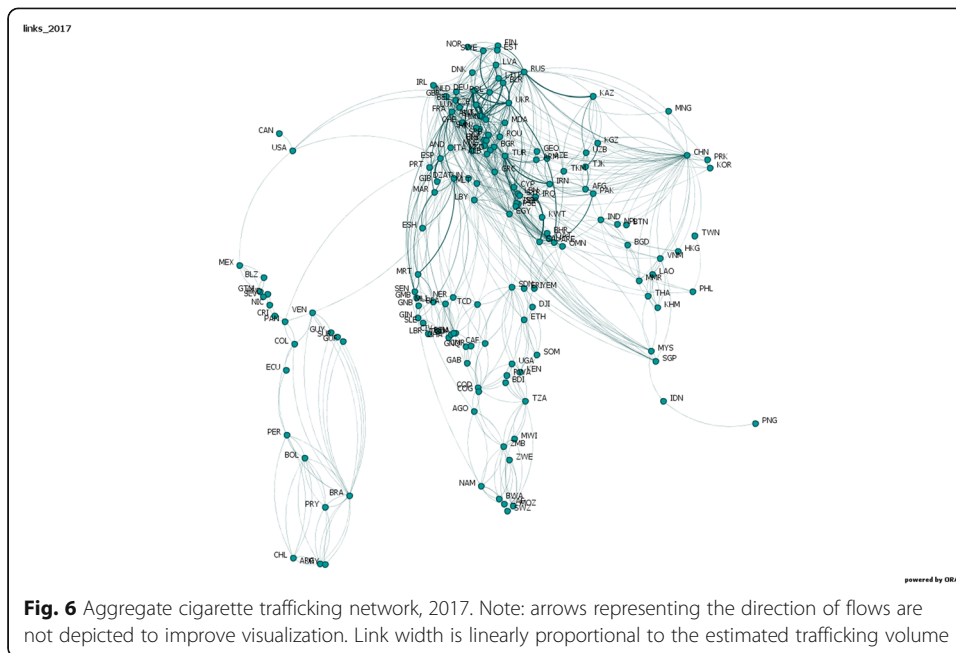
This result well explains why a methodology assuming a unidirectional flow between each pair of countries could not have been employed to map the cigarette trafficking network. The existence of bi-directional links is likely to be due to the intersection of geographically complex routes that often involve the shipment of different types of illicit cigarettes (counterfeits, illicit whites and contraband of genuine cigarettes with different origins). Indeed, countries involved in cigarette trafficking may be origin countries for a specific type of illicit cigarettes (e.g., contraband cigarettes) and destination countries for another type of illicit products (e.g., illicit whites originating in a neighboring country).

Even though the results highlight the presence of many cross-border illicit flows between the countries involved, the weighted density value of this network is fairly low, probably due to the high variability in the volumes of trafficked cigarettes.¹² However, network statistics must be interpreted taking into consideration that although the network includes 158 different countries, the analysis focuses on cigarettes destined for 57 of these countries (see the [Supplementary material](#) for the full list of included countries).

The temporal evolution of the statistics describing the network structure highlights a certain degree of stability in the overall structure of the network over time. In particular, the number of nodes and edges that constitute the network remains relatively constant over time. Moreover, while from a visual inspection some of the network measures in Table 4 appear to experience some temporary variations, most of these changes are in fact not statistically significant. This is the case, for example, of the (non-significant) increase in the network weighted density from 2012 to 2013, together with the increase in the two measures of network centralization (in-degree centralization and out-degree centralization).¹³ Visual inspection of Table 4 also suggests that the network assortativity (the extent to which nodes of similar degree are interconnected) follows an increasing trend until 2011, when it starts decreasing. This means that in the central years of the observation period countries appear to be involved in denser cigarette trafficking relationships, with countries of similar degree being clustered together. The increasing trend in network assortativity is correlated with

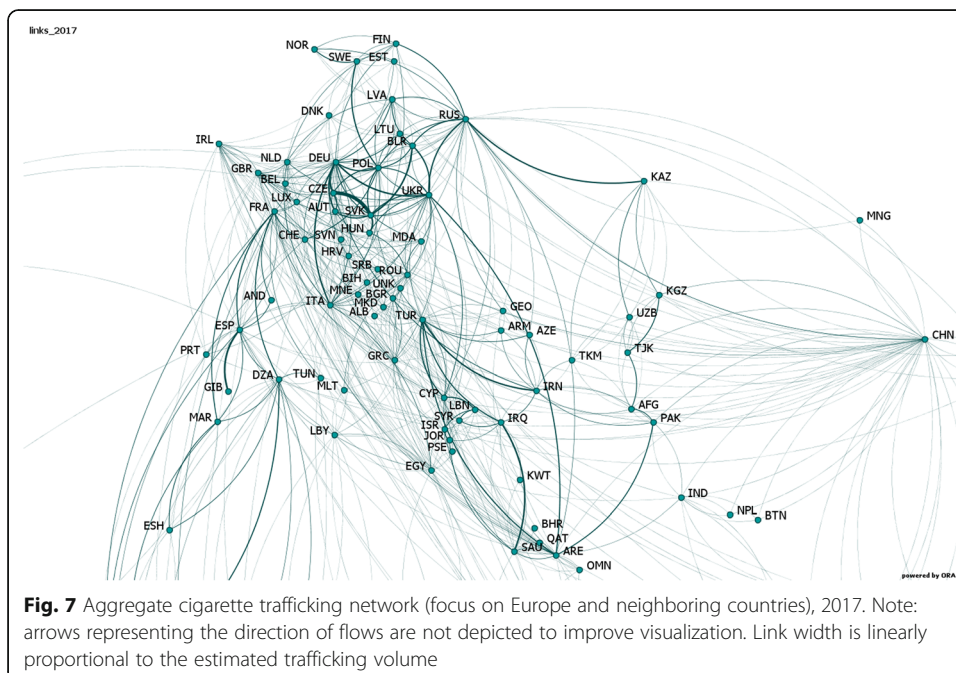
¹²The weighted density is computed as the ratio of the sum of the link values versus the maximum possible total sum for the network, this latter term being defined as the number of possible links times the largest link value (Altman et al. 2018).

¹³The statistical comparison of network measures for two different years is based on a t-statistic computed through bootstrapping techniques (see Altman et al. 2018, p. 1434). By relying on the bootstrapping technique, a small percentage of nodes and their attached links are removed from the network, the measure of interest is computed, and then a t-statistic and related standard errors are calculated. We performed 500 iterations of this perturbation method.



the increasing trend in the edge count, even though this last measure experiences smaller relative changes. A close inspection of results shows that newly added edges tend to connect to more peripheral regions of the network. Hence, the increase in network assortativity may be driven by the increased clustering of newly added nodes with low degree with already existing nodes of similar degree.

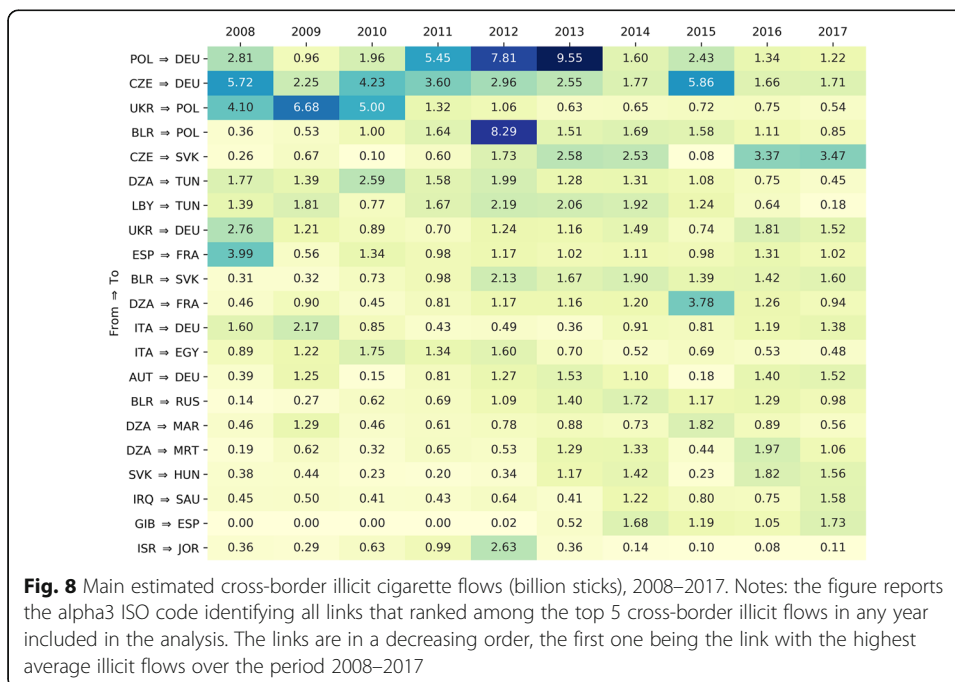
Despite the extensive structure of the trafficking network, illicit cigarettes do not move randomly across countries, and major flows tend to align along identifiable routes as postulated by opportunity theory (Felson and Clarke 1998). Results highlight the



predominant role in the network of countries located in Central and Eastern Europe, and the relevance of the North-Eastern Route (Aziani et al. 2019b) within the network, which ships illicit cigarettes originating in Eastern European countries into the European Union (see Fig. 8). The most substantial flows are related to routes bringing illicit cigarettes from origin countries to Germany, which is one of the main destination markets.

While the overall network structure remains fairly stable over time, as highlighted by the network statistics in Table 4 and discussed in the previous paragraphs (e.g., in terms of the number of existing edges), the spatial distribution of cross-border illicit cigarette flows (the edges' weights) experiences important changes. For instance, results highlight the emerging importance of Belarus as country of origin of illicit cigarettes. Conversely, the role of Ukraine as origin and transit country for illicit cigarettes attenuated in the period 2008–2017. The results also provide evidence of the presence of geographic displacement in how illicit cigarettes reach their destination markets. For example, the decrease in the volume of illicit cigarettes trafficked through some specific links along the routes with Germany as final destination (e.g., the “Czech Republic (CZE) → Germany (DEU)” connection) is compensated by the gain in importance of connections (e.g., the “Poland (POL) → Germany (DEU)” one) along alternative paths targeting the same final market. A close examination of the extent and of the drivers of these displacement mechanisms would make it possible to design policies that are able to effectively counter this crime instead of simply relocating trafficking routes geographically.

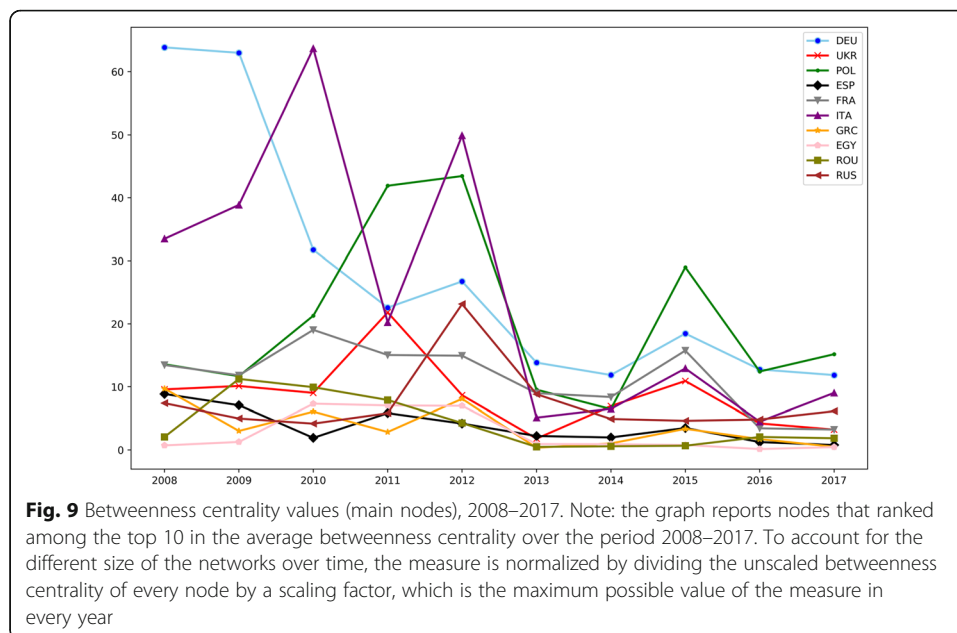
The reconstruction of the flows enables identification of the role that different countries play in the trafficking network (i.e., origin, transit or destination). Emerging results suggest that most countries have a multifaceted role, being for instance transit countries for illicit cigarettes of a given origin and destination markets for illicit cigarettes of another origin. As an example, considering countries that are relevant destination

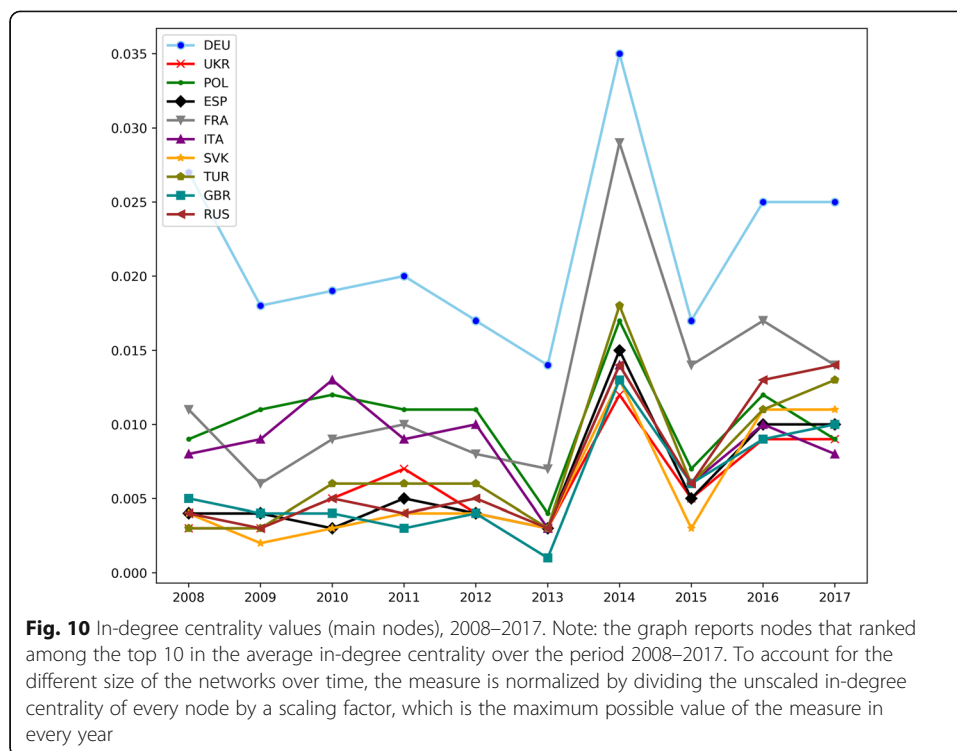


markets of illicit cigarettes, we note that some of them record high values of betweenness centrality (e.g., Germany, France, and Italy), suggesting they are also important transit hubs, while others report very low values (see Fig. 9). Remarkably, while Germany’s in-degree centrality is around the value of 0.025 both in 2008 and 2017 (despite experiencing some variations in between) (see Fig. 10), the betweenness centrality of Germany plummets from 2008 to 2017. This indicates that Germany lost over time its role as transit country and retained only a major importance as final destination market. The drop in the betweenness centrality value of Italy is also noticeable, and is in line with the findings of studies suggesting that the route bringing cigarettes from Southern to Western Europe across the Adriatic Sea is now much less exploited compared to few years ago (e.g., CSD 2018).

Discussion and conclusions

Scholars, policy makers, and law enforcement agencies interested in the empirical study and in combatting crime have constantly to face the severe lack of information characterizing criminal activities. This is especially the case of transnational crimes operated by criminal groups. The lack of key information on these crimes entails critically negative consequences for the possibility to design effective countering policies. This paper has made a first attempt to reconstruct, map and analyze a full transnational illicit cigarette trafficking network in a longitudinal perspective and using a cross-regional approach. Moving from previous studies focusing on the estimate of illicit cigarette consumption at the country level, it has presented an innovative methodology to map each step of the path followed by illicit cigarettes from the origin to the destination country and to measure the cross-border illicit cigarette flows between any pair of countries. By incorporating the volumes trafficked between each origin-destination country pair, we have been able to reconstruct the trafficking network of illicit cigarettes bound for consumption in 57 countries located in Europe, the Middle East, Central Asia, and North Africa. The results have highlighted the complexity of transnational cigarette trafficking





and its global dimension: almost 900 connections link the 158 countries involved in the trafficking of cigarettes targeting Europe, North Asia, the Middle East and Central Asia. However, not all trafficking connections are equally important, and while the illicit cigarette flows between two specific countries may be large (due to the fact that one or both countries have a key role as the origin, transit, or destination of illicit cigarettes), illicit cigarette flows between other neighboring countries may be negligible. The intersection of geographically complex routes, coupled with the different types of products being trafficked (e.g., counterfeits, illicit whites and contraband cigarettes with different origins), results in the bi-directionality of links in the network in the great majority of cases. The reconstruction of the flows also makes it possible to identify the roles that different countries play in the trafficking network, which in many cases appear to depend on the origin of illicit cigarettes flowing through the country.

The methodology presented does not come without limitations. Even though an extensive data collection on seizure cases has been implemented in order to obtain a representative sample of all seizures of illicit cigarette shipments, there are three main possible biases inherent to this data. First, due to language barriers, our data collection may not have resulted in a fully representative sample of all seizure cases reported by the news. Second, the media coverage of news on seizure cases is not uniform, because it depends on the country's sensitivity to the cigarette trafficking issue and on the type of seizure (media are more likely to report larger seizure cases). Third and perhaps most importantly, seizure cases are not an accurate reflection of the true routes used to traffic cigarettes, since the probability of law enforcement agencies detecting a shipment is not random and is correlated with the location and characteristics of the shipment itself.

To partially reduce the impact of the first of these biases, we searched for news items in multiple languages and from different news repositories. We attempted to control for the second bias by complementing seizure data retrieved from the news with information reported in press releases of custom authorities and official reports of different law enforcement and custom agencies. Finally, the third issue is inherent to many types of data on illicit activities. We partially mitigated this bias by inserting country-level fixed effects in the model estimating the weights of *'origin specific' likelihood networks*. More importantly, differently from previous attempts to map transnational illicit traffics (e.g., Boivin 2011, 2014; Berlusconi et al. 2017; Giommoni et al. 2017), the current study does not bind the identification of trafficking paths to the occurrence—and the size—of seizure cases. On the contrary, it exploits this important source of information together with detailed evidences of consumption of illicit cigarettes in the countries under analysis.

In addition to the possible biases related to seizure data, another limitation concerns the incorporation of geographical information to give better guidance to the algorithm determining the most plausible trafficking paths. Including a geographical element is essential in order to add a destination-specific variable to the procedure and incorporate a dimension of geographical convenience. These two elements are not embodied in consumption and seizure data exploited in step 1 “Building the weights of *'origin-specific' likelihood networks*”. However, geographical variables may not have been included in the determination of weights in the *'origin-specific' likelihood networks* since the impact of distance in predicting cigarette trafficking flows cannot be reliably estimated by relying on seizure data. In fact, the media disproportionally focus on news concerning large shipments that travel long distances by sea or air. While news reported by the media may be affected by different types of distortions when compared to the true picture of seizure cases, this bias may emerge as especially relevant during the data collection phase on seizure cases. Large shipments are more sensational and thus more appealing for the public; conversely, the general knowledge on cigarette trafficking is still too limited for the public to discriminate between different cigarette origins, which alleviates potential biases towards reporting news related to specific origin countries. To overcome this issue, we incorporated information on geographical convenience in a subsequent step by running Yen's algorithm in the likelihood networks of geographical distance and summing the likelihoods obtained with those derived from the *'origin-specific' likelihood networks*. The resulting estimates found support in qualitative evidence on trafficking routes emerging from the literature. Nonetheless, future studies should focus on integrating the current methodology with additional data or adjustments that allow the incorporation of geographic information through a more rigorous data-driven approach.

This study contributes to current knowledge in three main ways. First, the structure of the methodology presented could be further exploited to quantify and map other types of illicit flows, since it does not rely on any specific assumption concerning the illicit market on which we have focused. The core of the methodology consists of data on seizure cases and consumption of illicit goods which are already available for other types of illicit markets (e.g., drug trafficking, firearms trafficking). Conditional on data availability, the methodology could also be applied to estimate the cigarette trafficking network (or other trafficking networks) in a different area or at a different geographical

level (e.g., regional). Second, estimating and mapping illicit cigarette flows is the necessary first step to develop an inferential analysis for the identification of the risk factors of cigarette trafficking. The proposed methodology maps cross-border illicit cigarette flows by relying only on data that provide evidence of the existence of cigarette trafficking between two considered countries, thus making it possible to test for the effect of any determining factors on the size of these flows (e.g., economic incentives, government interventions). Finally, from a policy-making perspective, the analysis of the transnational cigarette trafficking network makes it possible to identify the borders with the highest risk and the different roles that countries play in this illicit market. Such analysis is crucial in order to plan effective policies to counter this crime.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1007/s41109-020-00265-3>.

Additional file 1: Table A1. Countries in the cigarette trafficking network.

Acknowledgements

The authors wish to thank Gian Maria Campedelli and Maria Jofre for their precious suggestions on earlier versions of this manuscript, Marina Mancuso and Federica Daphne Ierace for the support during the data collection phase, and Elias Mamo Fiedle for the technical help on the algorithmic specification.

Authors' contributions

CM, AA and MD have developed together the methodology for this study. CM and AA have conducted the quantitative analysis and written the paper. MD has provided inputs to the entire research process. All authors read and approved the final manuscript.

Funding

The authors have no support or funding to report.

Availability of data and materials

Seizure cases data collected for this study are available from the corresponding author on reasonable request. EPS and legal registered production data have been transmitted to the authors in the framework of NEXUS project (2019), which has received funding by PMI Impact, a global grant initiative by Philip Morris International to support projects dedicated to fighting illegal trade and related crimes. Restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Geographical information is available from CEPII at http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=6.

Competing interests

Part of the data analyzed in this study have been transmitted to the authors in the framework of NEXUS project (2019), which has received funding by PMI Impact, a global grant initiative by Philip Morris International to support projects dedicated to fighting illegal trade and related crimes. PMI Impact did not fund this paper, had no role in its writing, and did not exercise any editorial control.

Received: 25 October 2019 Accepted: 27 March 2020

Published online: 06 April 2020

References

- Akyeampong E (2005) Diaspora and drug trafficking in West Africa: a case study of Ghana. *Afr Aff* 104:429–447. <https://doi.org/10.1093/afaf/adi015>
- Allen E (2011) The illicit trade in tobacco products and how to tackle it. International Tax and Investment Center (ITIC), Washington
- Altman N, Carley KM, Reminga J (2018) ORA User's guide 2018. Center for the Computational Analysis of social and organizational systems. Carnegie Mellon University, Pittsburgh
- Alzaabi M, Taha K, Martin TA (2015) CISRI: a crime investigation system using the relative importance of information spreaders in networks depicting criminals communications. *IEEE Trans Inf Forensics Secur* 10:2196–2211. <https://doi.org/10.1109/TIFS.2015.2451073>
- Aziani A, Berlusconi G, Giommoni L (2019a) A quantitative application of Enterprise and social embeddedness theories to the transnational trafficking of cocaine in Europe. *Deviant Behav* 0:1–23. <https://doi.org/10.1080/01639625.2019.1666606>
- Aziani A, Dugato M (2019) ITTP NEXUS in Europe and beyond. *Transcrime* – Università Cattolica del Sacro Cuore, Milano
- Aziani A, Dugato M, Meneghini C (2020) A methodology for estimating the illicit consumption of cigarettes at the country level. *Glob Crime* 0:1–31. <https://doi.org/10.1080/17440572.2020.1724786>
- Aziani A, Kulick J, Norman N, Prieger JE (2017) Empty discarded pack data and the prevalence of illicit trade in cigarettes. Social Science Research Network, Rochester

- Aziani A, Mancuso M, Cinciripini L (2019b) North-eastern route. Mapping cigarette trafficking along the north-eastern route. In: Aziani A, Dugato M (eds) ITTP NEXUS in Europe and beyond. Transcrime – Università Cattolica del Sacro Cuore, Milano
- Berlusconi G, Aziani A, Giommoni L (2017) The determinants of heroin flows in Europe: a latent space approach. *Soc Netw* 51:104–117. <https://doi.org/10.1016/j.socnet.2017.03.012>
- Boivin R (2011) Le monde à l'envers: analyse de la structure du trafic transnational de drogues illicites Ph.D. thesis, Université de Montréal
- Boivin R (2014) Risks, prices, and positions: a social network analysis of illegal drug trafficking in the world-economy. *Int J Drug Policy* 25:235–243
- Bright D (2015) Identifying Key Actors in Drug Trafficking Networks. In: Bichler G, Malm A (eds) *Disrupting Criminal Networks: Network Analysis in Crime Prevention*. Lynne Rienner Publishers, pp 67–88
- Calderoni F (2014) A new method for estimating the illicit cigarette market at the subnational level and its application to Italy. *Glob Crime* 15:51–76. <https://doi.org/10.1080/17440572.2014.882777>
- Calderoni F, Dugato M, Aglietti V et al (2017) Price and non-price determinants of the illicit cigarette trade: analysis at the subnational level in the EU. In: Savona EU, Kleiman MAR, Calderoni F (eds) *Dual markets*. Springer, Cham, pp 267–286
- Campana P (2018) Out of Africa: the organization of migrant smuggling across the Mediterranean. *Eur J Criminol* 15:481–502. <https://doi.org/10.1177/1477370817749179>
- Chandra S, Joba J (2015) Transnational cocaine and heroin flow networks in western Europe: a comparison. *Int J Drug Policy* 26:772–780. <https://doi.org/10.1016/j.drugpo.2015.04.016>
- Chandra S, Yu Y-L, Bihani V (2017) How MDMA flows across the USA: evidence from price data. *Glob Crime* 18:122–139. <https://doi.org/10.1080/17440572.2016.1179632>
- Chionis D, Chalkia A (2016) Illicit tobacco trade in Greece: the rising share of illicit consumption during crisis. *Trends Organ Crime* 19:236–253. <https://doi.org/10.1007/s12117-016-9273-4>
- Cornish DB, Clarke RV (1987) Understanding crime displacement: an application of rational choice theory. *Criminology* 25: 933–948
- CSD (2018) The illicit trade of tobacco products along the Balkan route: Bulgaria, Greece, Italy and Romania. CSD, Sofia
- Decker SH, Townsend-Chapman M (2008) *Drug smugglers on drug smuggling: lessons from the inside*. Temple University Press, Philadelphia
- Desroches FJ (2007) Research on upper level drug trafficking: a review. *J Drug Issues* 37:827–844
- Dijkstra EW (1959) A note on two problems in connexion with graphs. *Numer Math* 1:269–271. <https://doi.org/10.1007/BF01386390>
- European Commission (2018) Excise duties: cigarettes - taxation and customs union - European Commission. *Eur. Comm. - Policies Inf. Serv* https://ec.europa.eu/taxation_customs/business/excise-duties-alcohol-tobacco-energy/excise-duties-tobacco/excise-duties-cigarettes_en. Accessed 7 May 2018
- Felson M, Clarke RV (1998) Opportunity makes the thief: practical theory for crime prevention. Home Office, London
- Ferrara E, De Meo P, Catanese S, Fiumara G (2014) Detecting criminal organizations in mobile phone networks. *Expert Syst Appl* 41:5733–5750. <https://doi.org/10.1016/j.eswa.2014.03.024>
- Fronzetti Colladon A, Remondi E (2017) Using social network analysis to prevent money laundering. *Expert Syst Appl Int J* 67: 49–58. <https://doi.org/10.1016/j.eswa.2016.09.029>
- Gallagher AWA, Evans-Reeves KA, Hatchard JL, Gilmore AB (2019) Tobacco industry data on illicit tobacco trade: a systematic review of existing assessments. *Tob Control* 28:334–345. <https://doi.org/10.1136/tobaccocontrol-2018-054295>
- Giommoni L, Aziani A, Berlusconi G (2017) How do illicit drugs move across countries? A network analysis of the heroin supply to Europe. *J Drug Issues* 47:217–240
- Guerette RT, Bowers KJ (2009) Assessing the extent of crime displacement and diffusion of benefits: a review of situational crime prevention evaluations*. *Criminology* 47:1331–1368. <https://doi.org/10.1111/j.1745-9125.2009.00177.x>
- Interpol (2014) *Countering illicit trade in tobacco products - a guide for policy-makers* Interpol
- Johnston N, Kegö W, Wenngren C (2016) *Cigarette smuggling: Poland to Sweden*. Institute for Security and Development Policy (ISDP), Sweden
- Joossens L, Chaloupka FJ, Merriman D, Yürekli AA (2000) Issues in the smuggling of tobacco products. In: Jha P, Chaloupka FJ (eds) *Tobacco control in developing countries*. Oxford University Press, New York, pp 393–406
- Joossens L, Raw M (1998) Cigarette smuggling in Europe: who really benefits? *Tob Control* 7:66–71. <https://doi.org/10.1136/tc.7.1.66>
- KPMG (2018) Project SUN. a study of the illicit cigarette market in the European Union, Norway and Switzerland 2017 Results. KPMG
- Kulick J, Prieger J, Kleiman MAR (2016) Unintended consequences of cigarette prohibition, regulation, and taxation. *Int J Law Crime Justice* 46:69–85
- Mackay J, Eriksen MP (2002) *The tobacco atlas*. World Health Organization, Geneva
- Mayer T, Zignago S (2011) Notes on CEPII's distances measures: the GeoDist database. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.1994531>
- Merriman D (2010) The micro-geography of tax avoidance: evidence from littered cigarette packs in Chicago. *Am Econ J Econ Policy* 2:61–84. <https://doi.org/10.1257/pol.2.2.61>
- OECD (2016) *Illicit trade: converging criminal networks*. OECD Publishing, Paris
- OECD, EUIPO (2017) *Mapping the real routes of trade in fake goods*. OECD Publishing, Paris
- Paoli L, Greenfield VA, Reuter P (2009) *The world heroin market: can supply be cut?* Oxford University Press, Oxford; New York
- Paoli L, Reuter PH (2008) Drug trafficking and ethnic minorities in Western Europe. *Eur J Criminol* 5:13–37
- Reuter PH (1983) *Disorganized crime: the economics of the visible hand*. MIT Press, Cambridge
- Reuter PH (2014) The mobility of drug trafficking. In: Collins J (ed) *Ending the drug wars*. Report of the LSE expert group on the economics of drug policy. The London School of Economics and Political Science, London, pp 33–40
- Reuter PH, Majmundar M (2015) *Understanding the U.S. illicit tobacco market: characteristics, policy context, and lessons from international experiences*. Committee on the illicit tobacco market: collection and analysis of the international experience; National Research Council, Washington

- Ross H, Vellios N, Smith KC et al (2015) A closer look at 'cheap white' cigarettes. *Tob Control*. <https://doi.org/10.1136/tobaccocontrol-2015-052540>
- Savona EU, Mancuso M (eds) (2017) Fighting illicit firearms trafficking routes and actors at European level. Final report of project FIRE. Transcrime - Università Cattolica del Sacro Cuore, Milano
- Shen A, Antonopoulos GA, von Lampe K (2010) The dragon breathes smoke. *Br J Criminol* 50:239–258. <https://doi.org/10.1093/bjc/azp069>
- Sin SS, Boyd MA (2016) Searching for the nuclear silk road: geospatial analysis of potential illicit radiological and nuclear material trafficking pathways. In: Volders B, Sauer T (eds) *Nuclear terrorism: countering the threat*. Routledge, New York
- Swami K, Judd CD, Orsini J (2009) Trace metals analysis of legal and counterfeit cigarette tobacco samples using inductively coupled plasma mass spectrometry and cold vapor atomic absorption spectrometry. *Spectrosc Lett* 42:479–490. <https://doi.org/10.1080/00387010903267799>
- Taha K, Yoo PD (2016) SIIMCO: a forensic investigation tool for identifying the influential members of a criminal organization. *IEEE Trans Inf Forensics Secur* 11:811–822. <https://doi.org/10.1109/TIFS.2015.2510826>
- Tinti P (2017) Migrant smuggling: paths from the horn of Africa to Yemen and Saudi Arabia Institute for Security Studies
- Transcrime (2016) An European outlook on the illicit trade in tobacco products. *Trends Organ Crime* 19:300–328. <https://doi.org/10.1007/s12117-015-9260-1>
- UNODC (2015) Drug money: the illicit proceeds of opiates trafficked on the Balkan route. United Nations Office on Drugs and Crime, Vienna
- von Lampe K (2002) The trafficking in untaxed cigarettes in Germany: a case study of the social embeddedness of illegal markets. In: *Upperworld and underworld in cross-border crime*. Wolf Legal Publishers, The Netherlands, pp 141–161
- von Lampe K (2011) The illegal cigarette trade. In: Natarajan M (ed) *International crime and justice*. Cambridge University, Cambridge, pp 148–154
- Webb EJ, Campbell DT, Schwartz RD, Sechrest L (1966) *Unobtrusive measures: nonreactive research in the social sciences*. Rand McNally, Oxford
- Yen JY (1971) Finding the K shortest Loopless paths in a network. *Manag Sci* 17:712–716
- Yürekli AA, Sayginsoy Ö (2010) Worldwide organized cigarette smuggling: an empirical analysis. *Appl Econ* 42:545–561. <https://doi.org/10.1080/00036840701720721>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
