UDK 339.194:663.974(100)

*Jože DAMIJAN, Črt KOSTEVC, Tjaša REDEK**

THE GRAVITY OF TOBACCO SMUGGLING. PREDICTING BILATERAL ILLICIT CIGARETTE FLOWS FROM AGGREGATE DATA

Abstract. *International tobacco smuggling remains an important concern for governments, tobacco manufacturers and health experts alike. While often linked to other forms of illegal activities, it also directly impacts government tax and health policies. Knowledge of factors that contribute to illicit tobacco trade and the existence of smuggling routes is strongly hampered by the lack of reliable data on bilateral flows of illicit tobacco. Therefore, reconstructing the trafficking routes and estimating the size of cross-border illicit flows are crucial steps for gaining better understanding of these crimes and enforcing actions aimed at countering them. This study is the first to use gravity estimation techniques to decompose aggregate illicit cigarette inflows for which data are available into their bilateral components. Our approach is a simple and effective method that can serve as a complement to other methods of pinpointing international trafficking flows such as empty discarded pack data or network analysis to help in the fight against illicit tobacco flows. Policymakers, customs officials as well as law enforcement can employ the presented methods as an additional tool in the fight against illicit trade.*

Keywords: *bilateral illicit cigarette trade, gravity model, predictive estimation*

Introduction

Illicit cigarette trade is an international criminal activity that causes significant losses in tax revenue, undermines the effectiveness of smoking reduction policies, and provides funding to organised crime and terrorist groups (Kulick et al., 2016). Research on cigarette smuggling is complicated

* *Jože Damijan, PhD, Professor, School of Business and Economics, University of Ljubljana, Ljubljana, Slovenia and KU Leuven, Leuven, Belgium; Črt Kosevc, PhD, Professor, School of Business and Economic, University of Ljubljana, Slovenia; Tjaša Redek, PhD, Professor, School of Business and Economics, University of Ljubljana, Ljubljana, Slovenia.*

DOI: 10.51936/tip.59.4.989–1009

by the lack of reliable data on both the illicit trade and the many factors that influence it. Criminals go to great lengths to conceal their activities, whereas users of illicit tobacco are unwilling to disclose information about their consumption. In addition, the dual demand for cigarettes (legal and illegal consumption) makes it difficult to identify illicit products. This not only limits the ability to analyse the patterns and dynamics of illicit tobacco trade, but also hinders policy efforts to combat it.

Research on illicit tobacco trade has generally focused on the prevalence of illicit tobacco products in specific markets, without focusing on the origins of contraband or trade patterns. While most of the empirical literature previously focused on cigarette smuggling into the major tobacco markets of the USA, the EU, the UK and Canada (Aziani et al., 2020, 2021; Baltagi and Levin, 1986, 1992; Merriman et al., 2000; Nicholson et al., 2014; Schwartz and Zhang, 2016), more recently the geographic scope and focus of the empirical literature has expanded (Europol, 2017; Meneghini et al., 2020; Metlzer and Martin, 2015) Meneghini et al. (2020) provide a very recent exception, applying network theory to redistribute the amount of illicit cigarettes consumed across identified paths in proportion to the estimated path-specific likelihood.

This study examines the drivers of illicit cigarette trade in Europe. Although assessments of the volume of illicit cigarette and tobacco trade are more widely available on the aggregate level, data on bilateral illicit flows are much more difficult to obtain. The objective of this study is to develop a method for estimating bilateral illicit cigarette trade between countries. We use the empirical approach developed by Proietti (2006) and Badinger and Cuaresma (2015) to decompose aggregate trade data into bilateral relationships using exogenous bilateral variables. The novelty of our approach compared to the existing literature is that: (i) to our knowledge, this is the first study to use a gravity-based decomposition of aggregate illicit cigarette inflows into bilateral components; (ii) this is one of the very few studies to estimate illicit trade with bilateral (origin-destination pairing) determinants; and (iii) one of the few studies based on panel data rather than 1-year cross-sections. The results presented in this article and the methodology employed in their estimation represent a valuable tool for policymakers and law enforcement alike in their fight against illicit tobacco smuggling. Moreover, the methodology could be extended and adapted to other areas of illicit trade as well as other illegal cross-border activities.

A key feature of bilateral illicit cigarette trade is the existence of smuggling routes between two countries. Whether for historical, logistical or other reasons, certain countries serve as either destination or transit markets for illicit cigarettes. This largely determines the volume of bilateral illicit cigarette trade between countries and, ultimately, the total inflow of such cigarettes into a market.

Other factors shown to contribute significantly to the volume of illicit trade are income differentials between the origin and destination markets. Combined with cigarette price disparities, income disparities have an influence from the perspective of both demand- and supply-side factors. On one hand, higher-income markets contain a potential pool of buyers for suppliers of illicit cigarettes while, on the other hand, conditions in lower-income markets encourage individuals and groups to engage in cigarette smuggling and illicit production despite the associated risks. Closely related to income inequality is price inequality. Even when the income differential is explicitly controlled, the price differential between the destination and origin countries serves as a robust predictor of illicit cigarette trafficking. The price differential acts as an indicator of potential profits for smuggling groups and motivates individuals and organisations to begin smuggling. While prices partly reflect living standards and purchasing power, they also reflect differences in tobacco tax policies across countries.

The remainder of the article proceeds as follows. The next section discusses the methodology and data and measurement issues. The third section outlines the main results, while the final one presents a conclusion.

Methods

As often occurs with economic applications, information on bilateral illicit cigarette trade relationships, which would be highly desirable from a policy perspective, is rarely available. While aggregate data on illicit cigarette and tobacco imports are more widely available, data on bilateral illicit flows are far more difficult to obtain. Still, data between two countries on a bilateral level are available for several variables of interest that could help determine the extent of the illicit cigarette trade and be used to explain variations in bilateral trade relations (Frankel and Romer, 1999).

As Proietti (2006) and Badinger and Cuaresma (2015) show, data on exogenous bilateral variables can be used to obtain reasonably good approximations of country-pair-specific bilateral data concerning variables of interest by decomposing aggregate data into bilateral relationships. Badinger and Cuaresma (2015) employ gravity trade models shown to predict bilateral trade levels on both the country and regional levels. They reveal that disaggregated bilateral flows between trading partners can be approximated when only aggregate destination country-specific data on a variable of interest are available. Their approach is based on Proietti (2006). We follow this approach to deconstruct aggregate illicit cigarette imports based on a combination of bilateral and unilateral determinants that account for the major push and pull factors in determining aggregate inflows of illicit cigarettes into destination markets. While we rely on the gravity framework as the basis of the estimation algorithm, other co-factors apply to cross-border tobacco flows and illegal trade activities. Details of the estimation algorithm are given in the appendix.

The illicit cigarette trade model is initially estimated on the aggregate level to explore the key destination country and bilateral determinants of the overall value of illegally imported cigarettes. The following specification is estimated:

$$\text{ill}_{it}^{cig} = \alpha + \beta_1 \text{bilateral controls}_{ijt} + \beta_2 \text{target market controls}_{it} + \beta_3 \text{country FE}_i + \beta_4 \text{time FE}_t + \varepsilon_{it} \quad (1)$$

where ill_{it}^{cig} is the estimated number of aggregate illicit cigarette inflows coming into target country *i* at time *t* (measured by the number of cigarette sticks).

Bilateral controls are as follows (with explanation provided)[1].

- Difference in GDP per capita between the destination country (*i*) and source country (*j*) at time *t* allows us to capture the difference in disposable income between countries as a measure of differences in demand for cigarettes. Disposable income has long since been used in models of cigarette demand to capture some of the non-price determinants of cigarette consumption (Baltagi and Levin, 1986; Hamilton, 1972; McGuinness and Cowling, 1975). Relative demand of the destination country serves as both a pull and a push factor for the influx of cigarettes from the origin country. We expect illicit cigarette trade to flow from poorer to richer countries.
- A key factor of illicit cigarette trade is differences in cigarette prices (of the brand most sold at official exchange rates in USD) between the destination (*i*) and source country (*j*) at time *t*. Increased prices of legal cigarettes stimulate the demand for illicit cigarettes while also creating the potential for a sizeable profit for would-be smugglers. As generally acknowledged in the literature (Meneghini et al., 2019), higher price differentials are likely to lead to bigger illicit cigarette flows between countries.
- Geographical proximity to source countries increases the likelihood of a destination country being involved in cigarette trafficking (Aziani and Dugato, 2019). To account for geographical proximity, we include distance between the origin (*i*) and destination markets (*j*) as a factor impacting the volume of cigarette trafficking. We expect distance to have a negative effect on illicit flows as it both increases the costs of transport and the likelihood of the illegal flow being interrupted by the authorities.

[1] *More details and definitions of variables are found in Table A1 in the Appendix.*

- Size of the market is captured by the population size of both the destination and origin markets. According to the gravity approach to international trade, large and similarly sized countries should trade more, while countries of very different size trade less. However, it often happens that relatively smaller countries act as a source for smuggled cigarettes into bigger more developed markets. The population of the target country along with the per capita income also capture the potential aggregate demand (Prieger and Kulick, 2018).
- The proportion of country aggregate income represented by activities hidden from authorities, also known as the shadow economy, accounts for both the propensity of the population to avoid paying taxes as well as society's acceptance of such behaviour (Meneghini et al., 2019)). By including both the extent of destination (*i*) and origin country (*j*) shadow economies at time *t,* we capture factors impacting both the demand and supply of illicit cigarettes. We assume that illicit flows of cigarettes come from countries with a high share of the informal economy and are similarly more likely to be smuggled into countries with a high share of informal economy.
- Apart from the distance between destination and origin countries, illicit trade is likely to be positively impacted by the existence of a shared border between countries (Calderoni et al., 2017). A contiguity indicator is included in the specification to capture the effect of a common border between countries (*i* and *j*). Since countries traditionally trade mostly with their neighbours, we expect illicit cigarette trade to be larger between neighbouring countries.
- The European illicit cigarette market is dominated by the existence of three crucial smuggling routes facilitating the supply of counterfeit tobacco products from the east, south and southeast (Aziani and Dugato, 2019; Meneghini et al., 2019). Indicators for the North-European, Balkan and Maghreb smuggling routes show cases when both origin and destination countries are found along one of the three main smuggling routes into Western Europe.
- Bilateral tourist flows measured by the logarithm of the number of tourists from the origin (j) to destination country (i) per year serve as an indication of the frequency of passenger travel between origin and destination countries and the potential for smuggling. Larger numbers of border crossings could be correlated with higher inflows of illicit goods.

Several destination- and origin-market-specific control variables were also used in the estimation model:

- The prevalence of other criminal activities in a country serves to indicate the overall propensity for crime as well as the social acceptability of (non-violent) crime. A few authors have noted that the target-market crime rate is a good predictor of illicit tobacco products (Melzer and Martin, 2016; Meneghini et al., 2019). Destination market crime rate as measured by the number of robberies per 1,000 people is included in the model. We expect the correlation between illicit cigarette trade and crime rates to be positive.

- Another key geographical feature that might influence the scale of illicit trade is sea access given that large transnational illicit cigarette flows often involve shipments by sea (Meneghini et al., 2019). Landlocked countries are hence less likely to be involved in illicit cigarette trade than countries with direct sea access.
- Finally, we intend to capture the licit productive and logistic capacity of cigarette manufacture and trade by indicating which countries produce cigarettes and/or have logistical centres for the distribution and sale of tobacco products. As illicit sales are often a by-product of legal production, production and logistical centres are included in the estimation. We expect that the presence of (legal) production and distribution centres in the destination will reduce illicit imports, whereas those in origin countries may increase them.
- Alongside the above, we explicitly control for other time-invariant country characteristics and temporal effects common to all countries by including additional country and time indicator variables.

Our analysis is based on data for 25 EU countries for the period 2008–2015 as destination markets for illicit cigarettes from the rest of the world. While destination countries are restricted to those in the EU, origin countries can be anywhere in the world (including the EU). Potentially, we could use 224 origin countries and 50,176 country pairs for every year in the sample, but data limitations (missing values) mean that some are not eligible. Aggregate illicit cigarette imports for the target markets are based on KMPG estimates (Project Stella). Although several methodological concerns have been raised over the years about the validity of the KPMG estimates, such as the lack of transparency and details about the data used in the model, overreliance on industry-produced data, risk of overestimation, and lack of external validation, they remain the best source of annual estimates on the prevalence of illicit tobacco trade in the EU (Calderoni, 2014b, 2014a). Data on illicit cigarettes are reported in numbers of cigarette sticks.

Cigarette prices in USD for the brand most sold at the official exchange rate come from the World Health Organisation and missing data for odd years were replaced with a linear interpolation. Information on gross domestic product and population size comes from the World Bank, while data on distance, contiguity, and sea access (landlocked) come from the

CEPII database. For bilateral distances, we use population-weighted distances in kilometres between a country's most important cities or agglomerations. The number of robberies per 1,000 people comes from the United Nations Office on Drugs and Crime (UNODC), whereas information on the extent of the shadow economy as a percentage of GDP is from Medina and Schneider (Medina and Schneider, 2018). Data on tourist arrivals were taken from the World Tourism Organisation, Yearbook of Tourism (2019).

Information on countries on the three main European cigarette smuggling routes was compiled by Meneghini et al. (2019), while the list of countries with cigarette production capacity and those with distribution facilities was provided by industry experts.[2]

Summary statistics for the key variables of interest are presented in Table 1.

Table 1: SUMMARY STATISTICS

Variable	Obs	Mean	Std. dev	Min	Max
Volume of illicit imports (cigarette sticks)	281,932	5245937	20500000	11088.9	264000000
Difference in GDP per capita (USD)	1,697,796	6.590	19.592	0.001	1035.630
Price difference USD	245,064	-0.021	3.762	-15.770	15.770
Share of shadow economy (destination) share	279,502	28.590	12.025	6.160	69.080
Share of shadow economy (origin) share	307,658	28.686	12.011	6.160	69.080
Population destination (in mil)	3,192,279	22.431	94.023	0.003	1392.730
Population origin (in mil)	3,191,560	22.447	94.026	0.003	1392.730
Bilateral distance in km	3,417,569	8477.661	4695.816	0.995	19888.660
Production centre origin country	3,688,102	0.165	0.371	0.000	1.000
Distribution centre destination country	3,688,102	0.222	0.416	0.000	1.000
Number of incoming tourists	8,628	249584.8	908152.2	1	1.41E+07
Contiguity indicator	3,411,968	0.012	0.110	0.000	1.000
Land locked country indicator	3,688,102	0.022	0.146	0.000	1.000
Number of robberies per thousand people (destination)	172,711	0.327	1.157	0.000	10.871

Source: Own calculations.

Our estimation approach requires us to estimate the gravity equation (1) for aggregate illegal imports into the destination market using a combination of bilateral (destination–origin pairings) and destination-specific variables. These estimates are used in the second stage to decompose aggregate illicit cigarette inflows into the (most likely) bilateral components. The first-stage estimates of equation 1 are presented in Table 2.

[2] *Other issues with the data and estimation are presented in the Appendix.*

Table 2: PANEL-DATA FEASIBLE GLS ESTIMATES OF (1) ON ILLICIT CIGARETTE TRADE INVOLVING EU COUNTRIES (2008–2015)

VARIABLES	(1) benchmark	(2) w/ routes	(3) w/ all
Log difference in GDP per capita	0.983*** (0.016)	0.982*** (0.016)	1.208*** (0.047)
Log price difference USD	0.036*** (0.010)	0.036*** (0.010)	0.074*** (0.025)
Log share of shadow economy (destination)	-0.697*** (0.021)	-0.695*** (0.021)	-0.575*** (0.068)
Log share of shadow economy (origin)	1.508*** (0.078)	1.506*** (0.078)	0.392 (0.429)
Log population size (destination)	0.874*** (0.007)	0.874*** (0.007)	0.800*** (0.026)
Log population size (origin)	-2.121*** (0.100)	-2.119*** (0.100)	-1.441 (1.853)
Production centre origin country	7.543*** (0.337)	7.544*** (0.337)	6.325 (5.691)
Production centre destination country	-0.668*** (0.012)	-0.667*** (0.012)	-0.294*** (0.045)
Distribution centre (origin)	-5.839*** (0.249)	-5.834*** (0.249)	-5.819 (3.701)
Distribution centre (destination)	1.093*** (0.022)	1.091*** (0.023)	1.128*** (0.075)
Log distance in km	-0.076*** (0.025)	-0.076*** (0.025)	0.066 (0.061)
Contiguity indicator	0.204*** (0.049)	0.204*** (0.049)	0.145* (0.083)
Land locked country indicator	-0.796*** (0.018)	-0.797*** (0.018)	-0.360*** (0.067)
Number of robberies per thousand people (destination)	0.086*** (0.021)	0.088*** (0.021)	0.786*** (0.090)
North European route indicator		-0.026 (0.048)	-0.156* (0.087)
Balkan route indicator		0.075 (0.047)	0.134 (0.105)
Maghreb route indicator		-0.114* (0.067)	
Log number of incoming tourists			0.033** (0.015)
Constant	16.855*** (0.506)	16.849*** (0.509)	18.160*** (5.159)
Origin country FE	YES	YES	YES
Time dummies	YES	YES	YES
N chi2 (df)	20,618 48227*** (162)	20,618 48246*** (165)	1,509 8616*** (147)

Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Source: Own calculations.

The results in Table 2 show that several robust correlations help determine the variance of aggregate illicit cigarette imports in our sample. Both income and cigarette price differences between destination and source markets are positively correlated with the volume of illicit cigarette trade. Yet, it is somewhat surprising that the share of the shadow (informal) economy in the destination country is negatively correlated with the total volume of illicit cigarette imports. This may reflect another aspect of a country's level of development not captured by GDP per person. More developed markets with stronger formal and informal institutions may have both a smaller share of the shadow economy and imported larger volumes of illicit cigarettes.

As expected, larger target markets attract bigger total inflows of illicit cigarettes. In contrast, smaller populations in source countries are associated with higher illicit flows. Legal production centres in the country of origin tend to be associated with higher illicit trade, while centres in the country of destination are associated with lower inflows. The results also show that overall one of the strongest associations, apart from legal production centres in the country of origin, is the negative correlation with distribution centres in the country of origin. Similarly, negative effects were also found for lack of access to the sea. When the number of tourist border crossings is included in the regression, a very strong correlation with illegal cigarette inflows is established. However, due to data availability issues, the sample size is greatly reduced, resulting in the loss of significance of several other coefficients.

While smuggling route indicators are generally not very accurately estimated and appear to have a relatively small impact on total illicit trade when other factors are controlled for, they tend to be associated with higher volumes of illicit trade for countries along the routes. Their inclusion is necessitated by the fact that they are consistently identified by industry experts as some of the key routes for supplying Western European countries with illicit tobacco. The three routes are relatively distinct in terms of both their source countries and the transition countries crossed *en route* to the destination. The fact that when all other controls are included the three main smuggling routes do not show a significant correlation with the aggregate number of imported cigarettes points to the strong cross-correlation of the smuggling routes with other regressors.

In terms of impact, the largest proportion of overall variance explained is seen in the share of the shadow economy in the country of origin, which accounts for nearly one-third of the variance in total illicit cigarette trade.

Results

Prediction of bilateral illicit cigarette flows

We use the specification (1) of the Badinger and Cuaresma (2015) algorithm to break aggregate country cigarette inflows down into estimates of bilateral flows. The regression-based decomposition matches the variation in aggregate flows of illicit cigarettes with the variation in the set of country-specific and bilateral covariates specified above. Based on this approach, one can construct predicted values for illicit cigarette trade on the bilateral level (between two trading partners) and decompose the total incoming volume of illicit cigarettes into predicted bilateral components.

Looking at the relative importance of the specific determinants of illicit cigarette inflows shown in Table 2, it is reasonable to assume that smuggled cigarettes are likely to originate from countries with a large share of the shadow economy (informal sector), a smaller population than in the target market, relatively low income (GDPpc) compared to the target market, lower cigarette prices than in the target market, and a relatively long distance to the target market. In addition, destination countries that lack local cigarette production centres, have high crime rates, and have access to the sea are more vulnerable to the influx of illicit cigarettes.

The biggest influence on the total volume of illicit cigarettes is the share of the shadow economy in GDP of the country of origin. Ukraine is one of the countries most often cited as a major source of illicit cigarettes (see Table 3 below). This is in no small part due to the relatively large share of the shadow economy in Ukraine (39.9% in 2014), which, combined with its relative proximity to the EU market, makes it a likely source of illicit cigarette trade. The share of the shadow economy is also high in Serbia, Belarus and Russia compared to several target markets in the EU. A 1%-reduction in the share of the shadow economy in countries exporting illicit cigarettes would reduce total illicit imports by almost 32% on average (holding other regressors constant).[3]

Predictably, the larger the market for illicit cigarettes (population size), the bigger the total volume of illicit cigarettes. Larger markets, especially those with higher purchasing power, are attractive target markets for cigarette smuggling. The difference in income (GDP per capita) between destination and origin markets is also a reliable predictor of illicit cigarette trafficking between these markets. The economic growth of less developed countries and their convergence with higher income markets would clearly

[3] *The estimate is based on the marginal effect of the Poisson-Pseudo Maximum Likelihood estimation of (1) (estimated at variable means) presented in Table A2 in the Appendix.*

limit the scope for illicit cigarette trade. A 1%-reduction in the income gap reduces the total volume of cigarette inflows by 11% on average.

Unlike total trade, distance appears to add to the potential for illicit cigarette trade. This could be an artifact of the existence of global cigarette production hotspots since cigarette production is very unevenly distributed geographically.

The presence of distribution centres in the destination country also appears to promote illicit cigarette trafficking through either existing logistical networks or demand-driven factors. The closure of logistics centres in the destination country reduces total predicted illicit imports by 8%. In contrast, the presence of production and distribution centres in the country of origin tends to reduce total illicit cigarette imports. In the case of distribution centres, the effect is quite substantial. The presence of distribution centres in the country of origin reduces total imports to destination countries by 21%.

Contrary to expectations, the share of the shadow economy in the destination country has a negative effect on total illicit cigarette imports when all other variables are controlled. This may be partly an outcome of the sample's composition since the largest quantities of illicit cigarettes go to the big developed markets of Western Europe, which also tend to have a smaller share of the shadow economy than in some Eastern and Central European countries.

Bilateral Decomposition of the Illicit Cigarette Trade

A breakdown of aggregate illicit cigarette imports for most European countries is shown in Table 3. Aggregate illicit cigarette inflows are decomposed based on the estimated bilateral decomposition (detailed in the Appendix) into bilateral flows of illicit cigarettes from the origin to the destination country. Based on the predicted bilateral illicit trade values between individual origin and destination countries, we calculate the share of the aggregate illicit cigarette imports accounted for by a particular trading partner and present the top origin countries by imported share in Table 3.

Ukraine is one of the top source countries, with 14 mentions as the first source, 2 mentions as the second source, and 1 additional mention. This is not unexpected given that several other studies highlight Ukraine as one of the main sources of illicit cigarettes entering Western Europe via the northeastern smuggling route (Aziani and Dugato, 2019; Meneghini et al., 2019). In 2017, the prices for premium cigarettes in Ukraine were 70% lower than in neighbouring EU countries, despite sharing land borders with the EU (Aziani and Dugato, 2019). Ukraine is followed by Serbia, Russia and Belarus as the most frequently cited countries of origin. Like with Ukraine, these countries also appear in other studies of illicit cigarette trafficking in the EU as countries of origin or transit along major smuggling routes (Melzer and Martin, 2016). Still, this does not mean that these countries always represent the actual origin of illicit cigarettes because other data sources would be needed to identify them.

Table 3: PRIMARY SOURCES FOR ILLEGAL CIGARETTE IMPORTS BY COUNTRY (TOP 5 SOURCE COUNTRIES AS PREDICTED BY THE EMPIRICAL MODEL)

Country	1st source	2nd source	3rd source	4th source	5th source
Austria	BIH (22.74%)	Greece (19.68%)	Turkey (19.20%)	Albania (15.35%)	Bulgaria (11.65%)
Belgium	Ukraine (34.21%)	Russia (24.53%)	Moldova (14.91%)	Poland (9.46%)	Estonia (8.75%)
Bulgaria	Ukraine (62.42%)	Serbia (37.58%)			
Czechia	Ukraine (41.28%)	Serbia (35.88%)	Belarus (12.69%)		
Croatia	Serbia (23.42%)	BIH (19.63%)	Greece (16.03%)	Turkey (15.58%)	Albania (11.38%)
Denmark	Romania (23.71%)	Bulgaria (22.71%)			
Estonia	Ukraine (49.25%)	Belarus (12.41%)	Russia (1.41%)		
Finland	Ukraine (46.38%)	Russia (22.07%)	Moldova (12.12%)	Estonia (7.44%)	Poland (6.88%)
France	Ukraine (12.22%)	Tunisia (9.61%)	BIH (9.30%)	Russia (8.79%)	Morocco (8.22%)
Germany	Ukraine (15.95%)	BIH (12.13%)	Russia (11.46%)	Greece (10.51%)	Turkey (10.26%)
Hungary	Turkey (53.92%)	Iran (18.47%)	BIH (14.22%)	Bulgaria (10.77%)	Ukraine (2.33%)
Ireland	Moldova (19.00%)	Romania (18.53%)	Poland (18.34%)		
Italy	Ukraine (12.18%)	Tunisia (9.59%)	BIH (9.27%)	Russia (8.77%)	Morocco (8.20%)
Latvia	Ukraine (40.10%)	Belarus (8.79%)	Russia (1.61%)		
Lithuania	Ukraine (46.87%)	Bulgaria (22.90%)	Russia (17.29%)		
Netherlands	Ukraine (47.47%)	Russia (23.31%)	Moldova (13.24%)	Poland (5.52%)	Estonia (5.33%)
Poland	Ukraine (44.97%)	Belarus (22.31%)	Russia (14.61%)		
Portugal	Tunisia (37.82%)	Algeria (35.18%)	Morocco (20.82%)	UAE (6.17%)	
Spain	Algeria (43.52%)	Ukraine (23.06%)	Tunisia (20.56%)	Bulgaria (18.39%)	Morocco (13.91%)
Slovakia	Ukraine (61.34%)	Belarus (24.07%)	Russia (14.59%)		
Slovenia	Serbia (20.82%)	BIH (18.09%)	Greece (15.54%)	Turkey (15.21%)	Albania (12.21%)
Sweden	Belarus (32.19%)	Ukraine (30.97%)	Russia (18.20%)	Poland (4.23%)	Lithuania (4.01%)
UK	Ukraine (23.99%)	Russia (17.11%)	Iran (11.85%)	Albania (9.68%)	

Note: Incomplete or lacking data meant that predicted flows could not be generated for incoming illicit trade into Cyprus, Malta and Romania.

Source: Own analysis.

Conclusions

The price structure of cigarettes makes them particularly attractive to smugglers who benefit from the fact that the production price of cigarettes represents only a small fraction of the retail price. The existence of smuggling weakens the effectiveness of taxes as the main component of tobacco control policy in three areas. First, smuggling and counterfeiting reduce the tax revenue that can be collected on legally sold cigarettes because, as tax rates increase, the volume of smuggled cigarettes is likely to increase. Second, attempts to control smuggling can tie up both private and law enforcement resources. Third, smuggling may provide opportunities for corruption that weaken the efforst of law enforcement. In addition, profits from cigarette smuggling can be used to fund other illegal and criminal activities (Melzer and Martin, 2016). Estimates of the extent and determinants of cigarette smuggling are therefore an important input for tobacco tax policy design. Since empirical research has focused mainly on the prevalence of illicit tobacco and its causes in large markets, the patterns of trade in illicit tobacco products have not been adequately explored in the literature.

This article has aimed to fill this gap in the literature by proposing an empirical approach to disaggregate aggregate illicit cigarette imports into bilateral imports from each source country. A gravity-based model of illicit tobacco trade was used as the basis for estimating the main elasticities of the bilateral variables, which in turn were used to construct the predicted bilateral components of aggregate illicit cigarette imports. Our estimates suggest that the most important determinants affecting bilateral illicit cigarette trade are income differences and cigarette prices between destination and source markets, the size of the destination market, the presence of (legal) production and distribution facilities in the origin/destination markets, and the size of the informal economy. Moreover, evidence suggests that cross-border movements of individuals may be a strong predictor of likely illicit flows.

The proposed approach could act as a practical complement to other methods of identifying the sources of illicit tobacco imports, such as the empty pack survey or expert opinion, and could help in developing a policy response to cigarette smuggling. With improved data availability in terms of both scale and country coverage, empirical models would be able to provide more accurate breakdowns of total inflows and help identify both likely smuggling routes and transportation modes of cigarette smuggling.

Estimating econometric models of total illicit imports into a target market inevitably encounters data limitations. Some of these limitations were already mentioned, but there are additional problems. First, illicit tobacco trade can be divided into contraband, counterfeit, illicit white tobacco, and unbranded tobacco. Each of these types of illicit tobacco products is potentially traded on its own terms. This should be reflected in the model and supported by reliable data for each type of illicit tobacco trade. Second, because cigarette smuggling is closely linked to other transnational activities of criminal networks, current data on human smuggling, drug smuggling, and other activities would improve the overall fit of the model. Finally, up-to-date information on the production and distribution capacity in each country would provide a solid basis for assessing local (legal) supply and allow conclusions to be drawn about the potential for illegally supplied tobacco products.

BIBLIOGRAPHY

Aziani, Alberto, F. Calderoni and Marco Dugato (2021): Explaining the Consumption of Illicit Cigarettes. Journal of Quantitative Criminology 37 (3): 751–789.

Aziani, Alberto and Marco Dugato (2019): ITTP NEXUS in Europe and Beyond. Transcrime – Università Cattolica del Sacro Cuore. Accessible at https://publires.unicatt.it/en/publications/ittp-nexus-in-europe-and-beyond, 15. 7. 2022.

Aziani, Alberto, Marco Dugato and Cecilia Meneghini (2020): A Methodology for Estimating the Illicit Consumption of Cigarettes at the Country Level. Global Crime 21 (2): 154–184.

Badinger, Harald and Jesus Crespo Cuaresma (2015): Aggregravity: Estimating Gravity Models from Aggregate Data. Applied Economics 47 (20): 2119–2126.

Baltagi, Bagi H. and Dan Levin (1986): Estimating Dynamic Demand for Cigarettes Using Panel Data: The Effects of Bootlegging, Taxation and Advertising Reconsidered. The Review of Economics and Statistics 68 (1): 148–155.

Baltagi, Bagi H. and Dan Levin (1992): Cigarette Taxation: Raising Revenues and Reducing Consumption. Structural Change and Economic Dynamics 3 (2): 321–335.

Calderoni, Francesco (2014a): A New Method for Estimating the Illicit Cigarette Market at the Subnational Level and Its Application to Italy. Global Crime 15 (1–2): 51–76.

Calderoni, Francesco (2014b): Mythical Numbers and the Proceeds of Organised Crime: Estimating Mafia Proceeds in Italy. Global Crime 15 (1–2): 138–163.

Calderoni, Francesco, Marco Dugato, Virginia Aglietti, Alberto Aziani and Martina Rotondi (2017): Price and Non-price Determinants of the Illicit Cigarette Trade: Analysis at the Subnational Level in the EU. Springer International Publishing.

Hamilton, James L. (1972): The Demand for Cigarettes: Advertising, the Health Scare, and the Cigarette Advertising Ban. The Review of Economics and Statistics 54 (4): 401–411.

Kulick, Jonathan, James Prieger and Mark A. R. Kleiman (2016): Unintended Consequences of Cigarette Prohibition, Regulation, and Taxation. International Journal of Law, Crime and Justice 46: 69–85.

McGuinness, Tony and Keith Cowlin, (1975): Advertising and the Aggregate Demand for Cigarettes. European Economic Review 6 (3): 311–328.

Medina, Friedrich and Leandro Schneider (2018): Shadow Economies Around the World: What Did We Learn Over the Last 20 Years? 18/17 IMF Working Paper Series, IMF. Accessible at https://www.imf.org/en/Publications/WP/Issues/2018/01/25/Shadow-Economies-Around-the-World-What-Did-We-Learn-Over-the-Last-20-Years-45583, 15. 7. 2022.

Melzer, Sharon and Chris Martin (2016): A Brief Overview of Illicit Trade in Tobacco Products. Accessible at https://www.oecd-ilibrary.org/content/component/9789264251847-8-en, 15. 7. 2022.

Meneghini, Cecilia, Alberto Aziani and Marco Dugato (2020): Modeling the Structure and Dynamics of Transnational Illicit Networks: An Application to Cigarette Trafficking. Applied Network Science 5 (1). Accessible at https://doi.org/10.1007/s41109-020-00265-3, 15. 7. 2022.

Meneghini, Cecilia, Marco Dugato and Alberto Aziani (2019): Connections—Explaining the Factors Driving Illicit Flows of Cigarette. In ITTP NEXUS in Europe and Beyond, 175–216. Università Cattolica del Sacro Cuore.

Merriman, David, Ayda Yurekli and Frank Chaloupka (2000): How Big is the Worldwide Cigarette Smuggling Problem? 1. Tobacco Control in Developing Countries.

Metlzer, Sharon and Chris Martin (2015): Illicit trade: Converging criminal networks: A brief overview of illicit trade in tobacco products. OECD Publishing. Accessible at https://read.oecd-ilibrary.org/governance/illicit-trade/a-brief-overview-of-illicit-trade-in-tobacco-products_9789264251847-8-en, 15. 7. 2022.

Nicholson, Andrew, Tracy M. Turner and Eduardo Alvarado (2014): Cigarette Taxes and Cross-border Revenue Effects: Evidence Using Retail Data. Public Finance Review 44 (3): 311–343.

Prieger, James E. and Jonathan Kulick (2018): Cigarette taxes and illicit trade in Europe. Economic Inquiry 56 (3): 1706–1723.

Proietti, Tommaso (2006): On the Estimation of Nonlinearly Aggregated Mixed Models. Journal of Computational and Graphical Statistics 15 (1): 18–38.

Schwartz, Robert and Bo Zhang (2016): Debunking the Taxation-contraband Tobacco Myth. CMAJ: Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne 188 (6): 401–402.

SOURCES

Europol (2017): 2017 Situation Report on Counterfeiting and Piracy in the European Union. Accessible at https://www.europol.europa.eu/publications-documents/2017-situation-report-counterfeiting-and-piracy-in-european-union, 15. 7. 2022.

APPENDIX

Table A1: VARIABLE DEFINITIONS

Variable	Definition
Aggregate illicit cigarette imports (dependent variable)	Logarithm of the number of sticks of illicit cigarettes estimated to have been imported into destination country *i* from the rest of the world at time *t*
Log difference in GDP per capita	Logarithm of the difference between GDP per capita of the destination country *i* and origin country *j* at time *t*
Log price difference (in USD)	Logarithm of the difference in cigarette prices (for a pack of the brand most sold at official exchange rates in USD) between the destination *i* and source country *j* at time *t*
Log share of shadow economy at the destination (origin)	Logarithm of the estimated share of shadow (underground) economy in GDP in the country of destination *i* (country of origin *j*) at time *t*
Log population size destination (origin)	Logarithm of the size of the population in the country of destination *i* (country of origin *j*) at time *t*
Production centre destination (origin)	Indicator (dichotomous) variable which assumes a value of 1 if the destination country i (country of origin *j*) had local cigarette production at time *t*
Distribution centre destination (origin)	Indicator (dichotomous) variable which assumes a value of 1 if the destination country i (country of origin *j*) had local cigarette distribution facilities at time *t*
Log distance in km	Logarithm of the population weighted distances in kilometres between country's most important cities or agglomerations between the country of origin (*j*) and destination country (*i*)
Contiguity indicator	Indicator (dichotomous) variable which assumes a value of 1 if destination country *i* and country of origin *j* share a common border and 0 otherwise
Landlocked country indicator	Indicator (dichotomous) variable which assumes a value of 1 if the destination country *i* is landlocked (has no direct sea access) and 0 otherwise
Number of robberies per 1,000 people (destination)	The number of robberies per 1,000 people in the country of destination country (*i*) at time *t*
North-European route indicator	Indicator (dichotomous) variable which assumes a value of 1 if the destination country (*i*) and origin country (*j*) are simultaneously located on the North-European smuggling route and 0 otherwise
Balkan route indicator	Indicator (dichotomous) variable which assumes a value of 1 if the destination country (*i*) and origin country (*j*) are both located on the Balkan smuggling route and 0 otherwise
Maghreb route indicator	Indicator (dichotomous) variable which assumes a value of 1 if the destination country (*i*) and origin country (*j*) are simultaneously located on the Maghreb smuggling route and 0 otherwise
Log number of incoming tourists (destination from origin)	logarithm of the number of tourists from the origin (*j*) to destination country (*i*) each year (*t*)

Source: Own analysis.

Data and estimation issues

Several issues limit the scope and practical applicability of the estimates. First, data limitations restrict the number of countries that can be used to generate predictions which, in turn, weakens the validity of the predictions of other bilateral flows. For instance, data for countries like Kosovo, Montenegro, FYR Macedonia, Albania, China, some of the Kavkaz republics etc. are either missing or at best partial, which prevents their inclusion in either aggregate estimation or the decomposition into bilateral flows. As these countries are commonly associated with cigarette manufacture and trade, they could play important roles as either countries of origin or transit for illicit flows of cigarettes. Their exclusion consequently also limits the overall validity of the model's predictions. For the highest predictive power to be achieved, it is imperative that data are collected for all relevant variables and countries. While partial coverage does yield some informative predictions, it cannot offer a complete picture of patterns and dynamics of illicit cigarette trade.

In addition to issues of data coverage, the remaining concern involves data quality and timeliness. Better quality estimates of overall illicit cigarette imports also lead to a more precise decomposition into bilateral flows.

All estimates provided in Table 3 are based on point estimates from a Poisson-Pseudo maximum likelihood regression of aggregate imports of illicit cigarettes and their respective confidence intervals. Many of the explanatory variables are highly correlated and their correlation impacts the estimated sign and size of the coefficient. Different measures of crime rates are highly interrelated as are income differences with differences in cigarette prices. On their own, the variables conform to theoretical and logical predictions while, when estimated together, the coefficients reflect their interplay.

Econometric framework

The starting point for econometric analysis is a linear model for (unobserved) bilateral data corresponding to *N* individuals,

$$y = X\beta + \varepsilon \tag{A1}$$

where y is an N^2-dimensional vector, X is a N^2 *x* *K*-dimensional known matrix of explanatory variables, ß is a *K x* 1-dimensional vector and is a N^2-dimensional error term, assumed to be $\varepsilon \sim N(0, \sigma^2 I_{N^2})$. Let the observed (aggregate) variable be given by an *N*-dimensional vector *Y* such that each element of the vector is given by $Y_i = \sum_{j=1}^{N} f(y_{ij})$ and *f(.)* is a twice continuously differentiable function.

Based on bilateral gravity model of trade in log form and the aggregation of the nonlinearly transformed bilateral variables, the model for the aggregated variable can be written as

$$Y = A(y) = (I_N \otimes i'_N) f(X\beta + \varepsilon) \tag{A2}$$

where i_N is a N-dimensional column vector of ones and $f(.)$ is a N^2-dimensional vector function where $\mathbf{f(y)}$ has a typical element given by $f(y_{ij})$. In a gravity context, with self-relationships of a country trading with itself ruled out, the number of observations would typically amount to $N(N-1)$.

While Badinger and Cuaresma (2015) consider two alternative estimators of (A1), (i) aggregate ML estimation of equation A2 and (ii) feasible GLS estimation of an augmented equation A2.

Aggregate maximum likelihood

The simpler of the two approaches is based on an interpretation of the aggregate model as being affected by shocks on the aggregate instead of the bilateral level. The true model can be thought of as being approximated by the specification

$$Y = (I_N \otimes i'_N) f(Xb) + \eta \tag{A3}$$

where $\eta \sim N(0, \sigma_\eta^2 I_N)$ is assumed. The normality assumption implies that the nonlinear least squares estimator of can be obtained as a solution to

$$\widehat{\beta_{ML}} = \max_{\beta} -\frac{N}{2}(Y - ((I_N \otimes i'_N) f(Xb))'(Y - (I_N \otimes i'_N) f(Xb) \tag{A4}$$

Linearised GLS

The setting given by (A2) corresponds to the case pf model of nonlinearly aggregated data which can be nested within the class of models investigated by Proietti (2006). An estimate of ß can be obtained using a linearised version of (A2). The Taylor expansion around some value of y is given by

$$Y \approx \overline{Y} + \Theta(y - \overline{y}) \tag{A5}$$

where $\Theta(x)$ is a N x N^2 Jacobian matrix of $A(x)$. Following Proietti (2006), one can use an iterative estimation method for *ß*. Starting with a trial value of $\tilde{y}$, the vector *ß* can be estimated using

$$\hat{\beta} = \left[(\Theta(\tilde{y})X)'\left(\Theta(\tilde{y})'\Theta(\tilde{y})\right)^{-1}(\Theta(\tilde{y})X)\right]^{-1}(\Theta(\tilde{y})X)'\left(\Theta(\tilde{y})'\Theta(\tilde{y})\right)^{-1}((\Theta(\tilde{y})\tilde{y} + \bar{Y} - Y) \tag{A6}$$

and the residuals on the bilateral level are given by:

$$\hat{\varepsilon} = \Theta(\tilde{y})'\big(\Theta(\tilde{y})\Theta(\tilde{y})\big)^{-1}(\Theta(\tilde{y})\tilde{y} + \bar{Y} - Y - \Theta(\tilde{y})X\hat{\beta}) \quad \text{(A7)}$$

The fitted values of the unobserved bilateral variable $\hat{y} = \tilde{X}\hat{\beta} + \hat{\varepsilon}$ are then used as the next trial value and the procedure is repeated until the change in the fitted bilateral variable is sufficiently small.

Poisson Pseudo-maximum likelihood estimation

Poisson pseudo-maximum likelihood estimators are very commonly used in the relevant literature on trade gravity and address several concerns associated with the use of other econometric techniques.

Consider the nonlinear form of the Anderson and Van Wincoop gravity model with a multiplicative error term (UNESCAP, 2020):

$$X_{ij}^k = \frac{Y_i^k E_j^k}{Y^k}\left(\frac{\tau_{ij}^k}{\Pi_i^k P_j^k}\right)^{1-\sigma_k} e_{ij}^k \quad \text{(A8)}$$

Writing (A8) in linearised form makes it clear that the error term is in logarithms too:

$$\log X_{ij}^k = \log Y_i^k + \log E_j^k - \log Y^k + (1-\sigma_k)\left[\log \tau_{ij}^k - \log \Pi_i^k - \log P_j^k\right] + \log e_{ij}^k \quad \text{(A9)}$$

The mean of e_{ij}^k depends on higher moments of e_{ij}^k, thus including its variance. If is e_{ij}^k heteroskedastic, which is highly probable in practice, then the expected value of the error term depends on one or more of the explanatory variables because it includes the variance term. This violates the first assumption of OLS and suggests that the estimator may be biased and inconsistent. It is important to note that this kind of heteroskedasticity cannot be dealt with by simply applying a robust covariance matrix estimator since it affects the parameter estimates in addition to the standard errors. The presence of heteroskedasticity under the assumption of a multiplicative error term in the original nonlinear gravity model specification requires the adoption of a completely different estimation methodology.

Santos Silva and Tenreyro (2006) present a simple way of dealing with this problem. They show that under weak assumptions – essentially just that the gravity model contains the correct set of explanatory variables – the Poisson pseudo-maximum likelihood estimator provides consistent estimates of the original nonlinear model. It is exactly equivalent to running a type of nonlinear least squares on the original equation. Since we are dealing with a pseudo-maximum likelihood estimator, it is not necessary that the data be in fact distributed as Poisson. Thus, although Poisson is more commonly used as an estimator for count data models, it is appropriate to apply it far more generally to nonlinear models such as gravity.

The Poisson estimator has a number of additional desirable properties for applied policy researchers using gravity models. First, it is consistent in the presence of fixed effects, which can be entered as dummy variables as in simple OLS. This is an unusual property of nonlinear maximum likelihood estimators, many of which have poorly understood properties in the presence of fixed effects. The point is a particularly important one for gravity modelling because most theory-consistent models require the inclusion of fixed effects by exporter and by importer.

Second, the Poisson estimator naturally includes observations for which the observed trade value is zero. Such observations are dropped from the OLS model because the logarithm of zero is undefined. However, they are relatively common in the trade matrix as not all countries trade all products with all partners (see e.g., Haveman and Hummels, 2004). Although the issue has mainly arisen to date in the context of goods trade, it is also relevant for services trade (see below). Dropping zero observations in the way that OLS does potentially leads to sample selection bias, which has become an important issue in recent empirical work (see below). Thus, the ability of Poisson to include zero observations naturally and without any additions to the basic model is highly desirable.

Third, interpretation of the coefficients from the Poisson model is straightforward and follows exactly the same pattern as under OLS. Although the dependent variable for the Poisson regression is specified as exports in levels rather than in logarithms, the coefficients of any independent variables entered in logarithms can still be interpreted as simple elasticities. The coefficients of independent variables entered in levels are interpreted as semi-elasticities, asunder OLS.

In Table A2, we present estimates of the gravity equation (1) on aggregate illicit cigarette inflows (KPMG data) using PPML.

Table A2: RESULTS OF PPML ESTIMATION OF EQUATION 1 ON YEAR 2010 (DEPENDENT VARIABLE: LOGARITHM OF KPMG ESTIMATES OF THE AGGREGATE ILLICIT CIGARETTE INFLOWS)

	coefficient	se(b)
Log share of shadow economy (origin)	0.318	0.016
Log population size (destination)	0.122	0.001
Log difference in GDP per capita	0.109	0.002
Indicator distribution centres (destination)	0.077	0.002
Log distance in km	0.028	0.010

	coefficient	se(b)
Log price difference USD	0.004	0.001
Number of robberies per 1,000 people (destination)	0.001	0.000
Contiguity indicator (origin destination country)	0.000	0.000
Balkan route indicator	0.000	0.000
North European route indicator	0.000	0.000
Maghreb route indicator	0.000	0.000
Number of serious assaults per thousand (destination)	-0.010	0.000
Landlocked country indicator	-0.012	0.000
Production centre destination country	-0.020	0.000
Production centre origin country	-0.110	0.005
Log share of shadow economy (destination)	-0.133	0.003
Log population size (origin)	-0.153	0.007
Distribution centre (origin)	-0.215	0.009
Destination country dummy	YES	
Origin country dummy	YES	
# of observations	2520	

Note: Marginal effects (at means) reported

Source: Own calculations based on data collected.