

Cigarette Taxes and Illicit Trade in Europe

Appendix

James E. Prieger and Jonathan Kulick

This appendix contains more complete information about the data and additional regressions to which the main text refers.

A. Are taxes lowered in response to IRTC?

As discussed in the text in section IV.B.1, reverse causality between IRTC and cigarette taxes may cause downward bias in the estimated price coefficient. Since there are claims in the literature that some countries have responded to rising illicit trade by lowering cigarette taxes (Joossens & Raw, 2000)—albeit not for our countries and period—we investigate this further here. First, there are few examples of cigarette taxes falling by any appreciable amount in the data, even though IRTC has been rising, particularly in Western Europe (see Figure 1 in the main text). Figure 8 shows the trends in variable *ExTaxHypo* by country. The only countries with nontrivial decreases in this cigarette tax index are Bulgaria, Romania, and the U.K. Examination of the co-movements in illicit market share and *ExTaxHypo* in these countries in Figure 9 shows that only in Bulgaria does the tax index fall when the illicit share rises (and in only one year, 2010).

We repeat estimation IV 7.2 without these three countries. The results in Table 16 show that the price coefficient is indeed higher (and still highly significant) with the smaller sample. However, given the relatively small difference in the two price coefficients compared to their standard errors, there is no clear evidence against the consistency of the main estimates from IV 7.2.

B. Construction of the hypothetical excise taxes

In section IV.B.1 of the text, a hypothetical excise tax is used as an instrument for cigarette prices. Here we describe how that variable, *ExTaxHypo* (or T^h as it is denoted here), is constructed. We begin with the following identities:

$$P^h = P_0^h + T^h \quad (\text{A-1})$$

$$T^h = T_{as} + T_{av}^h \quad (\text{A-2})$$

In the expression, h superscripts are for hypothetical quantities. P^h is the tax-included retail sales price (TIRSP) with the hypothetical tax, P_0^h is the pre-excise-tax base price to be used in constructing the hypothetical tax, and T^h is the hypothetical overall excise tax (i.e., the sum of the specific and ad valorem excise taxes). In equation (A-2), T_{as} is the actual amount-specific excise tax (taken to be exogenous, thus also used for the hypothetical tax) and T_{av}^h is the hypothetical ad valorem excise tax amount (not the rate). The actual ad valorem tax amount cannot be used since it was calculated from a base that was presumed to be endogenous.

To calculate P^h , define α to be the ad valorem excise tax rate as a percentage of TIRSP. Rate α is taken to be constant for a country and year, even though the actual base from which the ad valorem tax is calculated may differ from country to country. Treating the ad valorem tax as a fraction of TIRSP may appear odd, since TIRSP itself includes excise taxes, but we do this for two reasons. It matches how the tax data are presented in both the Euromonitor and the alternative European Commission data (discussed in a later section of the appendix), as a fraction of TIRSP. Also, EU law requires minimum

taxes on cigarettes,⁴⁹ and the minimums are stated as a fraction of TIRSP. Thus, regardless of the actual rate and base each country chooses, policymakers must calculate and consider their taxes in terms of a fraction of TIRSP.

Then, from (A-1), (A-2), and by definition of α , we have

$$P^h = P_0^h + T_{as} + \alpha P^h \quad (\text{A-3})$$

or, after rearranging terms,

$$P^h = \frac{P_0^h + T_{as}}{1 - \alpha} \quad (\text{A-4})$$

Combining (A-1) and (A-4) and rearranging terms yields the final expression for $ExTaxHypo_{it}$:

$$T^h = \frac{P_0^h + T_{as}}{1 - \alpha} - P_0^h \quad (\text{A-5})$$

Matching to the data described in the text, T_{as} is $SpecTax_{it}$ and α is $AVtax_{it}$, both given in the Euromonitor data. The base price P_0^h , which is required to be plausibly exogenous since $ExTaxHypo_{it}$ will be used as an instrument, remains to be chosen. Define P_0^a to be the actual TIRSP less the overall actual excise tax (i.e., $P_0^a = CigPrice_{it} - ExTax_{it}$). Then we choose P_0^h for all observations to be the average value of P_0^a in the sample. Since the result does not vary over time, it cannot be endogenous due to time-varying factors affecting actual prices. Since P_0^h does not vary across countries, it cannot be correlated with endogenous country-specific factors. Instrument $ExTaxHypo_{it}$ varies over time and country because both T_{as} and α vary over both.

C. A consumption gap analysis of illicit trade

To complement the main estimations using the Euromonitor estimates of IRTC, we also calculated our own estimates of illicit market share. The consumption gap analysis in this section

⁴⁹ See Council of the EU's Directive 2011/64/EU of 21 June 2011 on the structure and rates of excise duty applied to manufactured tobacco.

provides further evidence that higher licit cigarette prices have sizeable and statistically significant impacts on illicit market share, even when avoiding industry- and third-party estimates of IRTC. Following the approach suggested by Blecher (2010), we compare survey estimates of cigarette consumption within each country with licit sales and ascribe gaps between the two to consumption of illicitly obtained cigarettes. We begin with the identity that the quantity consumed in a country and year must equal the quantity supplied from all sources.⁵⁰ Sources of supply include Q^S , domestic licit sales (including domestic production sold in-country, legally imported and taxed foreign product sold in-country, and domestic sales purchased by foreign visitors). Another source of supply is Q^{FL} , foreign legal product that is brought into the market legally by consumers (such as during a cross-border trip). Domestic supply is reduced by X^{FL} , domestic licit sales taken out of country by visitors. The final source of supply is Q^{CC} , counterfeit and contraband illicit product (whether produced domestically or from abroad). If there is outflow of domestically produced illicit product, it is netted out of Q^{CC} . Defining net foreign legal supply as $N^{FL} = Q^{FL} - X^{FL}$, the identity between domestic consumption Q^D and supply is thus

$$Q^D = Q^S + N^{FL} + Q^{CC} \quad (\text{A-6})$$

By definition, Q^D equals S , the number of smokers in the country, times A , the average smoking intensity. While S can be reasonably well estimated from survey data, it is well known that A is likely to be underreported (as \tilde{A}) by survey respondents (Warner, 1978; Merriman, 2000). Assume that underreporting is by a constant multiple $0 < \theta < 1$, so that estimated consumption ignoring the misreporting is $\tilde{Q}^D = \tilde{A}S$ but that actual consumption is

$$Q^D = \tilde{Q}^D / \theta \quad (\text{A-7})$$

Rearranging the terms in equation (A-6) yields

⁵⁰ We thus set aside issues involving inventories, since cigarettes are perishable product.

$$Q^{CC} = Q^D - Q^S - N^{FL}$$

and so the illicit market share Q^{CC}/Q^D , denoted M^{CC} , is

$$M^{CC} = \frac{Q^D - Q^S - N^{FL}}{Q^D} = 1 - \frac{Q^S + N^{FL}}{\tilde{Q}^D} \theta \quad (\text{A-8})$$

Assume a linear fixed-effects regression model for M_{CC} , so that

$$M_{it}^{CC} = \alpha_i + \beta' x_{it} + u_{it} \quad (\text{A-9})$$

Equating the right sides of equations (A-8) and (A-9), we have:

$$1 - \theta \frac{Q^S + N^{FL}}{\tilde{Q}^D} = \alpha_i + \beta' x_{it} + u_{it} \quad (\text{A-10})$$

$$\frac{Q_{it}^S}{\tilde{Q}_{it}^D} = \frac{1 - \alpha_i}{\theta} - \left(\frac{\beta}{\theta}\right)' x_{it} - \left(\frac{u_{it}}{\theta_i} + \frac{N^{FL}}{\tilde{Q}^D}\right) \quad (\text{A-11})$$

Define y_{it} to be $Q_{it}^S/\tilde{Q}_{it}^D$ and let \ddot{a} represent a variable a that has been demeaned by the within

transformation: $\ddot{a}_{it} = a_{it} - \frac{1}{T} \sum_{t=1}^T a_{it}$. Then equation (A-10) after transformation becomes:

$$\dot{y}_{it} = -\left(\frac{\beta}{\theta}\right)' \dot{x}_{it} + v_{it} \quad (\text{A-12})$$

where the new error term v_{it} is a mean-zero function of the terms in the final parentheses on the right side of equation (A-11) for all periods.

If the original error u was strictly exogenous in equation (A-9), and there is either no net foreign legal supply ($N^{FL} = 0$) or it is also strictly exogenous, then v is exogenous in equation (A-12). However, if cigarette price is an element of x , it is likely correlated with v through the term N^{FL}/\tilde{Q}^D . When licit prices rises within the country, then N^{FL} may rise (as consumers obtain more cigarettes abroad and foreigners buy fewer cigarettes within the country) and \tilde{Q}^D may fall (since demand is responsive to price). Since N^{FL}/\tilde{Q}^D enters v negatively, price and the error term are therefore negatively correlated, and we thus expect there to be downward bias on the estimated coefficient $-\beta_p/\theta$ on price in equation (A-11). The bias thus would exaggerate the estimate of β_p , the causal impact of price on actual IRTC.

Conceptually, the bias occurs because part of the impact on observed “illicit trade”, as defined by the method above, will actually be increasing net foreign legal supply.

There are three potential responses to this potential bias. The first would be to ignore it, as appears to have been done in previous literature (e.g., Blecher, 2010). The second would be to gather data on net foreign legal supply, so that N^{FL} could be moved out of the error term v . Since the only consistently calculated estimates of which we are aware for the EU during our entire period are those from KPMG in the Project Star reports, we do not follow this approach, since we wish to avoid using KPMG data where possible in this estimation. The third response, which we adopt, is to recognize the issue but argue that the results are illustrative nonetheless. The KPMG data indicate that foreign legal supply from border-crossing by consumers happens, as shown in Figure 10, where it is labeled “non-domestic legal”. However, the figure also shows that it is a minor part of overall trade, more stable than illicit trade, and less responsive to price changes.⁵¹ During 2009-2011, for example, prices sharply increased but non-domestic legal consumption did not change. We thus anticipate (but cannot prove) that any bias is limited.

Estimation of regression equation (A-7) yields estimates of coefficient vector β/θ , but the individual elements of β are identified only to scale. Estimation of equation (A-7) by itself therefore can show whether price has a statistically significant effect on illicit trade share but cannot reveal the magnitude of the effect. If data on domestic consumption from KPMG are allowed into the estimation, then scalar θ can be identified. KPMG estimates domestic consumption for the EU in its Project Star reports. The largest part of consumption is from legal domestic sales, which are readily observed from industry and tax data. KPMG then adjusts the figures to arrive at total consumption by adjusting for non-

⁵¹ The standard deviation of the non-domestic legal consumption figures in the graph is 0.6, whereas for illicit consumption it is 1.1.

domestic consumption as estimated from empty discarded pack studies.⁵² While the final consumption estimate relies on calculations and estimates by KPMG, the figures should be less contentious than the firm's direct estimates of IRTC. Treating the KPMG consumption data as Q^D , equation (A-2) in logs can be treated as a second regression equation to estimate θ .

Data for S and \tilde{A} are taken from Eurobarometer (various years) surveys. Not all years are available.⁵³ Furthermore, in some years, \tilde{A} was not directly reported because survey responses were instead quantized (i.e., instead of mean cigarettes per day reported, the fraction of smokers falling into various consumption ranges was given). For such years, mean smoking intensity was estimated from the quantal data by fitting a lognormal distribution to the data via MLE and calculating the implied mean based on the results.⁵⁴ With the resulting estimates of \tilde{A} and S , estimates of \tilde{Q}^D were formed as their product. Comparison of these estimates of \tilde{Q}^D with the KPMG estimates of Q^D is in Figure 11. The scatterplot shows that there is very high correlation between the two estimates (0.98). The slope of the line of best fit for the logged data, from estimation OLS 17.1 in the first column of Table 17, is close to one, as it must be if equation (A-2) is correct.⁵⁵ The implied value of θ from the regression is 0.63.⁵⁶ Thus, for every 10 cigarettes apparently actually consumed, smokers claim on average to have smoked only 6.3.

The results of estimation of equation (A-7), with and without using KPMG data on domestic consumption, are in second and following columns of Table 17. The dependent variable y in the first

⁵² Empty pack surveys are based on a large sample of packs collected via formal sampling plans in various cities throughout the countries. Once packs are collected, they are examined to determine the proportion of packs that did not originate domestically.

⁵³ Eurobarometer surveys are available covering data from 1995, 2002, 2005, 2006, 2008, 2009, 2012, and 2014.

⁵⁴ Each country was allowed to have its own set of location and scale parameters for the lognormal distribution.

⁵⁵ The estimated slope of the line is 1.03, with 95% confidence interval spanning one [0.961,1.097].

⁵⁶ Given the log form of equation (A-2), θ is the exponentiated constant from the log-log regression.

equation is constructed using domestic licit sales data from Euromonitor for Q^S . Estimation WI-OLS 17.1 (WI for within-transformed data) returns estimates of β/θ from OLS estimation on the pooled, demeaned data. The estimate-to-scale of the impact of cigarette price on illicit share is positive and statistically significant at the 5% level. Estimation WI-OLS 17.2 is similar except that year fixed effects are accounted for; the estimate-to-scale of the impact of price on illicit share is now significant at the 1% level. These results bolster the main conclusions of the paper, namely that price increases spur illicit trade, without indicating the magnitude of the marginal effect. However, if the previous estimate of θ (from estimation OLS 17.1) is employed, the implied estimates of β (the elasticity of illicit market share with respect to licit cigarette price) from estimations WI-OLS 17.1 and 17.2 are 0.30 and 0.47, respectively. These estimates are higher than the elasticities found from the fixed-effects estimates reported in Table 7 (perhaps due to the bias discussed above).

The final two columns of Table 17 contains estimates of β and θ separately, estimated from nonlinear seemingly unrelated regression (NLSUR) estimation of equation (A-7) and the log form of equation (A-2).⁵⁷ Joint estimation can improve the precision of the estimates, as well as identifying β . Estimation NLSUR 17.1 does not include year fixed effects, while NLSUR 17.2 does. The implied elasticities of illicit market share with respect to price are 0.41 and 0.56. Again, these estimates are higher than those from Table 7. The estimates of θ are quite close to those from estimation OLS 17.1, and are significant at the 5% level.⁵⁸ In summary, the work here using the consumption gap analysis corroborates the finding in the paper that increasing the licit price of cigarettes (e.g., through taxation) has sizeable and statistically significant impacts on IRTC.

⁵⁷ Estimation is with the `nlsur` command in Stata 14.1.

⁵⁸ The transformation from the estimated $\ln(\theta)$ to the reported θ is accomplished with the `nlcom` command.

D. Alternative data for prices and taxes

The European Commission produces annual reports on cigarettes prices and taxes.⁵⁹ Through 2010, the reported price is for the most popular price category (MPPC; i.e., the average price of the modal type of cigarette sold, usually king-size filter brands). After that year, weighted-average price (WAP; i.e., average revenue for cigarettes) is reported. The variables are converted to real terms using the same method as for all nominal data. Inclusion of year fixed effects in the estimations mitigates any impacts of this discontinuity in the time series. Tax data are also reported, as in the Euromonitor data: ad valorem and specific excise taxes, along with the VAT rate (expressed as a percentage of tax-included retail sales price, TIRSP). There is a high degree of concordance between the European Commission and the Euromonitor data. Log prices from the two sources exhibit correlation of 0.97. The instruments *ExTaxHypo* from equation (1) constructed from the EC and EM sources has correlation of 0.92. The VAT rates from the two sources have correlation of 0.97. Given these high correlations, the similarity of the new results in Table 12 with those in Table 7 is not surprising.

E. Alternative measures of corruption and governance

The estimations in Table 13 and Table 14 employ alternative regressors to *NotCorrupt*. Here these alternatives are described. Two measures from Transparency International are employed.⁶⁰ The first is the Corruption Perceptions Index, for which higher values indicate less corruption. Before a methodological change in 2012, the index was valid for cross-sectional comparison of countries but not across years. Therefore in estimations OLS 13.1 and FE 14.1 using this variable, year fixed effects are included to remove the changes over time due solely to the noncomparability of the data across years.

⁵⁹ See Part III (Manufactured Tobacco) of the excise duty tables of the yearly *Excise duties and transport, environment and energy taxes* reports published by the European Commission Directorate General for the Taxation and Customs Union.

⁶⁰ Data are from transparency.org, from which the Excel spreadsheet “cpi 1995_2013.xls” was obtained.

The other variable from the same source is the country's cross-sectional rank of its Corruption Perceptions Index. The ranks are rescaled to the unit interval to account for differing number of countries in different years. This variable is another attempt to account for the inherent noncomparability of the index across years. This variable is used in estimations OLS 13.2 and FE 14.2.

The World Bank dataset from which *NotCorrupt* was drawn, World Governance Indicators, also contains indices of the rule of law and the effectiveness of government in the country. The former measure reflects perceptions of the extent to which "agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence."⁶¹ The latter measures "perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies."⁶² For both indices a higher score indicates stronger governance. Years 1999 and 2001 are unavailable and were linearly interpolated. These data are constructed to be appropriate for use in panel data, and are comparable in the cross-section and the time-series. The rule of law variable is used in estimations OLS 13.3 and FE 14.3. Similar estimations with the government effectiveness index yielded nearly identical results (which is unsurprisingly, given the 95% correlation between the two regressors) and therefore are not shown in the tables.

Data from the United Nations on total police personnel at the national level per 100,000 people are taken from UN Office on Drugs and Crime (UNODC).⁶³ The counts include personnel in public

⁶¹ See info.worldbank.org/governance/wgi/rl.pdf.

⁶² While the government effectiveness index may appear to be less germane to IRTC than the other corruption and governance measures, we included it because Melzer (2010) found it to be significant in some of her cross-sectional regressions of illicit cigarette consumption share.

⁶³ See unodc.org/unodc/en/data-and-analysis/United-Nations-Surveys-on-Crime-Trends-and-the-Operations-of-Criminal-Justice-Systems.html.

agencies whose principal functions are the prevention, detection and investigation of crime and the apprehension of alleged offenders, excluding administrative support staff. Since there were many missing values, linear interpolation was used to fill in the gaps in the data. Estimations OLS 13.4 and FE 14.4 use the linearly interpolated data, while estimations OLS 13.5 and FE 14.5 use only the actual data.

Additional References

The following references are cited only in the appendix.

- Merriman, David. (2001). "Understand, measure, and combat tobacco smuggling." In *World Bank economics of tobacco toolkit* (Yurekli, A., and Beyer, J. de, eds.). World Bank Working Paper 80580.
- Eurobarometer. (2003). "Smoking and the environment: Actions and attitudes." *Special Eurobarometer, 183*, Wave 58.2 - European Opinion Research Group EEIG. Brussels: European Commission.
- Eurobarometer. (2006). "Attitudes of Europeans towards tobacco." *Special Eurobarometer, 239*, Waves 64.1 – 64.3 – TNS Opinion & Social. Brussels: European Commission.
- Eurobarometer. (2007). "Attitudes of Europeans towards tobacco." *Special Eurobarometer, 272C*, Wave 66.2 – TNS Opinion & Social. Brussels: European Commission.
- Eurobarometer. (2009). "Survey on tobacco: Analytical Report." *Flash Eurobarometer, 253*, *The Gallup Organisation*. Brussels: European Commission.
- Eurobarometer. (2010). "Tobacco." *Special Eurobarometer, 332*, Wave 72.3 – TNS Opinion & Social. Brussels: European Commission.
- Eurobarometer. (2012). "Attitudes of Europeans towards tobacco." *Special Eurobarometer, 385*, Wave EB77.1 – TNS Opinion & Social. Brussels: European Commission.
- Eurobarometer. (2012). "Attitudes of Europeans towards tobacco and electronic cigarettes." *Special Eurobarometer, 429*, Wave 82.4 – TNS Opinion & Social. Brussels: European Commission.
- Melzer, Sharon A. (2010). *Counterfeit and contraband cigarette smuggling: Opportunities, actors, and guardianship* (Unpublished doctoral dissertation). American University, Washington, D.C.
- Warner, K. E. (1978). "Possible increases in the underreporting of cigarette consumption." *Journal of the American Statistical Association, 73*(362):314-318.

Appendix Tables

Table 16: Fixed-effects panel regressions of IRTC share on cigarette prices (reduced sample)

Y = illicit market share	IV 16.1
Log(real cigarette price)	0.458 (0.173)***
Freedom from Corruption (WGI)	-0.028 (0.050)
Income (GNI) per capita (€1,000)	-0.014 (0.013)
GNI per capita squared	-0.000 (0.000)
1 st stage <i>F</i> statistic on excluded instruments	64.8
Sargan-Hansen statistic (<i>p</i> -value)	0.480
<i>N</i>	215

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; SE's (in parentheses) account for clustering by country.

Note: The instruments are *LaborTax*, *VAT*, and *ExTaxHypo*. Country and year fixed effects are included.

Table 17: Estimations for the consumption gap analysis

	OLS 17.1	WI-OLS 17.1	WI-OLS 17.2	NLSUR 17.1	NLSUR 17.2
	θ	β/θ	β/θ	β	β
<i>Equation (A-7)</i>					
Log(real cigarette price)		0.481 (0.184)**	0.745 (0.215)***	0.406 (0.200)**	0.558 (0.272)**
Income (GNI) per capita (€1,000)		0.017 (0.018)	0.023 (0.028)	0.019 (0.013)	0.021 (0.016)
GNI per capita squared		-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)*	-0.000 (0.000)*
<i>Equation (A-2) in log form</i>					
Coefficient on \tilde{Q}^D	1.029 (0.033)***			1.028 (0.032)***	1.028 (0.033)***
Log(θ)	-0.460 (0.308)			-0.476 (0.302)	-0.482 (0.311)
θ	0.631 (0.194) [†] *			0.621 (0.188) [†] **	0.617 (0.192) [†] **
Within-transformed data	N	Y	Y	Y	Y
Year fixed effects	N	N	Y	N	Y
<i>N</i>	98	103	103	74	74

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; SE's (in parentheses) account for clustering by country.

[†] Significance stars for θ are for the null hypothesis that $\theta = 1$.

Note: Estimates from the WI-OLS estimation are for β/θ from equation (A-7), and thus are not comparable to the direct estimates of β from the NLSUR estimations. The within-transformation and year fixed effects are for equation (A-7) only, when included.

Appendix Figures

Figure 8: The cigarette tax index (ExTaxHypo)



Figure 9: Illicit market share and ExTaxHypo in Bulgaria, Romania, and the U.K.

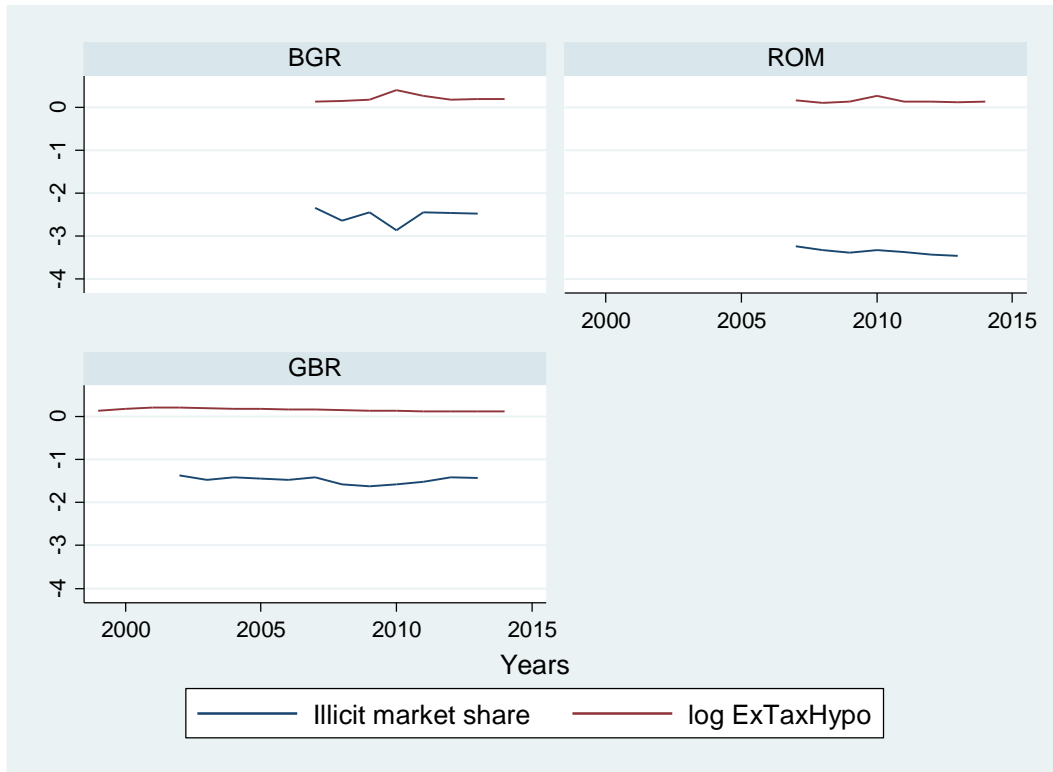
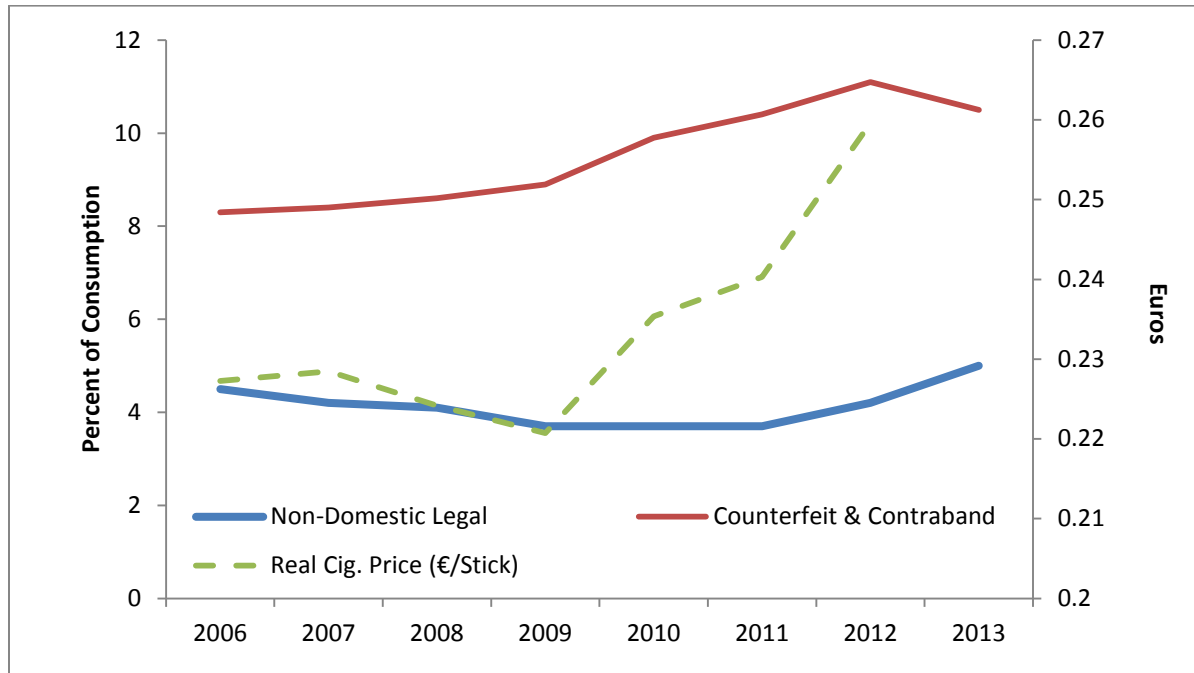


Figure 10: Data from KPMG Project Star for cigarette consumption and prices in the EU



Note: Price are national averages weighted by population. The data source is KPMG Project Star reports, various years.

Figure 11: Stated cigarette consumption vs. estimated actual consumption

