Annex B. Methodology

The main producing economies and key transit points for counterfeit and pirated goods were identified for each of the selected industries: foodstuff; pharmaceuticals; perfumery and cosmetics; leather articles and handbags; clothing and textile fabrics; footwear; jewellery; electronics and electrical equipment; optical, photographic and medical apparatus; toys and games.

For each product category the exercise was? carried out in several steps:

1) Economies were ranked according to their propensity to be an economy of provenance for counterfeit goods in trade in this product. The resulting index is called GTRIC-e.

2) An indicator of the relative comparative advantage for producing a given good was calculated for each economy (RCAP-e). This is the first "filter" to be used in the analysis.

3) For each economy an indicator of the relative comparative advantage for being a transit point in global trade in a given good was calculated (RCAT-e). This is the second "filter" to be used in the analysis.

4) Both filters (RCAP-e and RCAT-e indicators) were applied for every economy with a high GTRIC-e score. This indicates whether the given economy is a producing one, or a potential transit point.

5) Some additional descriptive statistical analysis checked the modes of transport and the size of shipments on the selected trade routes.

It should be highlighted that the framework presented below relies on a set of methodological assumptions. For transparency purposes all are spelt out in the text.

Construction of GTRIC-e for each product category

For each product category the first step was to rank all the known provenance economies by their relative intensity of exporting fakes. This distinguished the key transit points in trade with fake goods in a given product category. Each of these key points then was further investigated to determine its exact role in trade in fakes in the analysed sector.

The most intense provenance economies were identified using an index that ranked them according to their relative propensity to be an economy of provenance for counterfeit goods (GTRIC-e). The index is based on the data on global customs seizures and data on imports. It takes into account 1) the absolute value of exports of fakes from a given economy (in USD); and 2) the share of fakes in total exports in a given product category from a given economy.

The construction of GTRIC-e directly relied on the methodology introduced in the OECD-EUIPO (2016) study. A detailed description of the methodology used to calculate the GTRIC-e is provided below.

Importantly, two assumptions are made to calculate the GTRIC vectors. The first is that the volume of seizures of a given product or from a given source economy is positively correlated with the actual intensity of trade in counterfeit and pirated goods in that product category or from that economy. The second assumption acknowledges that this relationship is not linear, as there might be some biases in the detection and seizure procedures. For instance, the fact that infringing goods are detected more frequently in certain categories could imply that differences in counterfeiting factors across products merely reflect that some goods are easier to detect than others, or that some goods, for one reason or another, have been specially targeted for inspection.

Within each product category, GTRIC-e was constructed in four steps:

1) For each reporting economy, the seizure percentages for provenance economies were calculated.
2) For each provenance economy, aggregate seizure percentages were formed, taking the reporting economies' share of sensitive imports as weights.
3) From these, each economy's counterfeit source factor was established, based on the provenance economies' weight in terms of global trade.
4) Based on these factors, the GTRIC-e was formed.

Step 1: Measuring reporter-specific seizure intensities from each provenance economy

v_{epi} is economy *i*'s registered seizures of all types of infringing goods included in a given product category *p* that originate from economy *e* at a given year in terms of value.

γ_{epi} is economy *i*'s relative seizure intensity (seizure percentage) of all infringing items within the product category that originate from economy *e*, in a given year:

$$\gamma_{epi} = \frac{v_{epi}}{\sum_e v_{epi}}, \text{ such that } \sum_e \gamma_{epi} = 1 \quad \forall\, i$$

Step 2: Measuring general seizure intensities of each provenance economy

The general seizure intensity for economy *e* within the product category *p*, denoted Γ_{ep}, is then determined by averaging seizure intensities, γ_{epi}, weighted by the reporting economy's share of world imports from known counterfeit and pirate origins.[1] Hence:

$$\Gamma_{ep} = \sum_i \varpi_{pi} \gamma_{epi}$$

where the weight of reporting economy *i* is given by

$$\varpi_{pi} = \frac{m_{epi}}{\sum_i m_{epi}}$$

with m_{epi} is economy *i*'s imports of goods in a given product category *p* from economy *e* at a given year in terms of value, so that $\sum_i \varpi_{pi} = 1 \ \forall p$

Step 3: Measuring partner-specific counterfeiting factors

$m_{ep} = \sum_i m_{epi}$ is defined as the total registered world imports of all sensitive goods in the product category p from provenance economy *e*.

$m_p = \sum_e m_{ep}$ is defined as the total registered world imports of all sensitive goods in the product category *p* from all provenance economies.

The share of provenance economy *e* in world imports of all sensitive goods in the product category *p*, denoted s_{ep}, is then given by:

$$s_{ep} = \frac{m_{ep}}{m_p}, \text{ such that } \sum_e s_{ep} = 1, \ \forall p$$

From this, the economy-specific counterfeiting factor is established by dividing the general seizure intensity for economy *e* with the share of world imports from *e* within the product category *p*:

$$CF_{ep} = \frac{\Gamma_{ep}}{s_{ep}}$$

Step 4: Establishing GTRIC-e

Gauging the magnitude of counterfeiting and piracy from a provenance economy perspective can be done in a similar fashion as for sensitive goods. Hence, a general trade-related index of counterfeiting for economies (GTRIC-e) is established along similar lines and assumptions:

- The first assumption (A3) is that the intensity by which any counterfeit or pirated article from a particular economy is detected and seized by customs is positively correlated with the actual amount of counterfeit and pirate articles imported from that location.
- The second assumption (A4) acknowledges that assumption A3 may not be entirely correct. For instance, a high seizure intensity of counterfeit or pirated articles from a particular provenance economy could be an indication that the provenance economy is part of a customs profiling scheme, or that it is specially targeted for investigation by customs. The importance that provenance economies with low seizure intensities play regarding actual counterfeiting and piracy activity could therefore be under-represented by the index and lead to an underestimation of the scale of counterfeiting and piracy.

As with the product-specific index, GTRIC-e is established by applying a positive monotonic transformation of the counterfeiting factor index for provenance economies using natural logarithms. This follows from assumption A3 (positive correlation between seizure intensities and actual infringement activities) and assumption A4 (lower

intensities tend to underestimate actual activities). Considering the possibilities of outliers at both ends of the GTRIC-e distribution – i.e. some economies may be wrongly measured as being particularly susceptible sources of counterfeit and pirated imports, and vice versa – GTRIC-e is approximated by a left-truncated normal distribution as it does not take values below zero.

The transformed general counterfeiting factor across provenance economies on which GTRIC-e is based is therefore given by applying logarithms onto economy-specific general counterfeit factors (see, for example, Verbeek, 2000):

$$cf_{ep} = \ln(CF_{ep} + 1)$$

In addition, it is assumed that GTRIC-e follows a truncated normal distribution with $cf_{ep} \geq 0$. Following Hald (1952), the density function of the left-truncated normal distribution for cf_{ep} is given by

$$g_{LTN}(cf_{ep}) = \begin{cases} 0 & if\ cf_{ep} \leq 0 \\ \dfrac{g(cf_{ep})}{\int_0^{\infty} g(cf_{ep})\partial cf_{ep}} & if\ cf_{ep} \geq 0 \end{cases}$$

Where $g(cf_{ep})$ is the non-truncated normal distribution for cf_{ep} specified as:

$$g(cf_{ep}) = \frac{1}{\sqrt{2\pi\sigma_{cf}^2}} \exp\left(-\frac{1}{2}\left(\frac{cf_{ep} - \mu_{cf}}{\sigma_{cf}} \right)^2 \right)$$

The mean and variance of the normal distribution, here denoted μ_{cf} and σ_{cf}^2, are estimated over the transformed counterfeiting factor index, cf_{ep}, and given by $\widehat{\mu}_{cf}$ and $\widehat{\sigma}_{cf}^2$.

This enables the calculation of the counterfeit import propensity index within each product category *p* (GTRIC-e) across provenance economies, corresponding to the cumulative distribution function of cf_{ep}.

Construction of RCAP-e and RCAT-e

Relative comparative advantage for production of a given good (RCAP-e)

The first statistical filter that can be used to tell producers from transit points looks at the production capacities of a given economy in a given sector. The rationale behind this test is simple: production activity often relies on certain skills, or resources. It also exhibits certain returns to scale properties that results in specialisation of this particular economy in the production of that good. Hence, production of counterfeits in a sector is

more likely to occur in a known provenance economy that specialises in the legitimate production of a given good, than in a country without production capacity in a given sector.

This specialisation of a given trading economy in production of a given good is captured by an indicator of the relative comparative advantage for production (RCAP-e). The indicator looks at the share of industrial activity in a given sector with the total industrial activity in a given economy.

Construction of this indicator is based on industry statistics. Importantly, these statistics are based on a different taxonomy than the trade statistics, hence a matching exercise was performed (see Box B.1). A detailed description of the methodology used to calculate the RCAP-e is provided below.

Box B.1. Product classification methods

Although the datasets on trade and industrial activity in principle classify the same goods, they differ in the taxonomies used. Industry data (output) are extracted from the industrial statistics database of the United Nations Industrial Development Organization (UNIDO). These data are classified according to the categories of industrial activity (ISIC-Rev3) at a two-digit level. Trade data and seizure data are classified using the Harmonized Tariff Schedule (HTS) classification scheme. These differences are due to the fact that although they cover the same issues, they were created and are run independently.

In order to create the RCAP-e indicator, the HS code that refers to the GTRIC-p tables and to categories of international trade are matched with the relevant categories of industrial activity (ISIC). This is done following the concordance tables proposed by the United Nations Statistics Division (available at: http://unstats.un.org/unsd/cr/registry/regot.asp?Lg=1).

More formally, the revealed comparative advantage in production for an economy *e* in a given product category *p* (RCAP-e) measures whether this economy produces more of this given type of product as a share of its total production than the "average" country:

$$RCAP_{ep} = \frac{y_{ep} \Big/ \sum_p y_{ep}}{\sum_e y_{ep} \Big/ \sum_e \sum_p y_{ep}}$$

where y_{ep} is the output of product *p* by economy *e* in a given year.

Relative comparative advantage for being a transit point (RCAT-e).

The relative comparative advantage for being a transit point in global trade (RCAT-e) is the second filter used to determine the actual role of a provenance economy. This indicator represents the degree to which a given economy specialises in re-exporting a given product, e.g. through development of advanced logistical infrastructure, or by its convenient geographical location. Consequently, it is assumed that such factors that facilitate transiting of genuine products will also facilitate transit of fake products in the same product categories.

The RCAT-e indicator is calculated by comparing relative volumes of re-export of a given good to the shares calculated for other exporting economies. This is done based on

re-export data that come from the UN Comtrade database. A detailed description of the methodology used to calculate the RCAT-e is provided in Annex B.

Formally, the revealed comparative advantage in transit for an economy *e* within a given product category *p* (RCAP-e) measures whether this economy re-exports more goods of this given type of product as a share of its total re-exports than the "average" country:

$$RCAT_{ep} = \frac{x_{ep} \Big/ \sum_p x_{ep}}{\sum_e x_{ep} \Big/ \sum_e \sum_p x_{ep}}$$

where x_{ep} is re-exports of product *p* by economy *e* in a given year.

Application of both filters

Once the statistical filters (RCAP-e and RCAT-e indicators) are constructed, they are applied to distinguish the producing economies from the key potential transit points. Both filters are applied for every economy on the top provenance list for counterfeit goods, i.e. economies with a high GTRIC-e score. The selection of top economies is done arbitrarily, depending on the distribution of the GTRIC within a given product category.

The rationale for using the filters is as follows: if an economy is *not* a significant producer of a fake good (i.e. its RCAP-e for this good is low) and/or is a large re-exporter of this good in legitimate trade (i.e its RCAT-e for this good is high), then it is likely to be a transit point.

On the other hand, if this top listed provenance economy of counterfeit goods within the product category is a significant producer (i.e. has a high RCAP-e score) or is a small re-exporter (i.e. has a low RCAT-e score), it is likely to be a producer of the fake goods.

This exercise results in a list of producers and a list of transit points. Together with the information on the place of seizure, this will allow the development of maps of trade in fake goods in given product categories, showing key producers, main transit point and main destination points.