New Estimates of Trade Mis-invoicing and Methods based on artificial intelligence

26 June 2018

Kathy Nicolaou-Manias
Armand Desjardines

UNCTAD UNECA UNODC

Overview of the Presentation

Presentation Overview

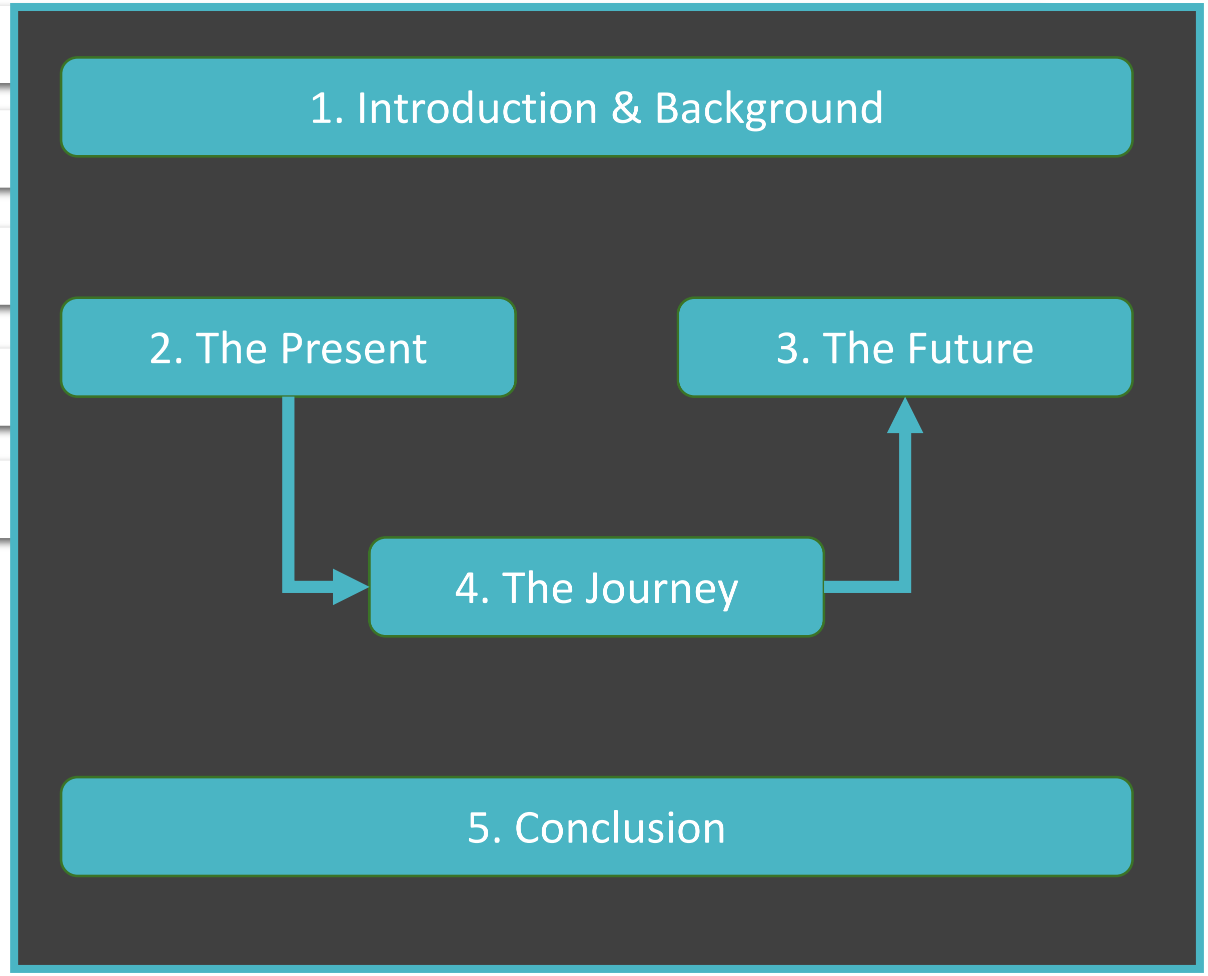
1 Introduction

2 The Present

3 The Future

4 The Journey

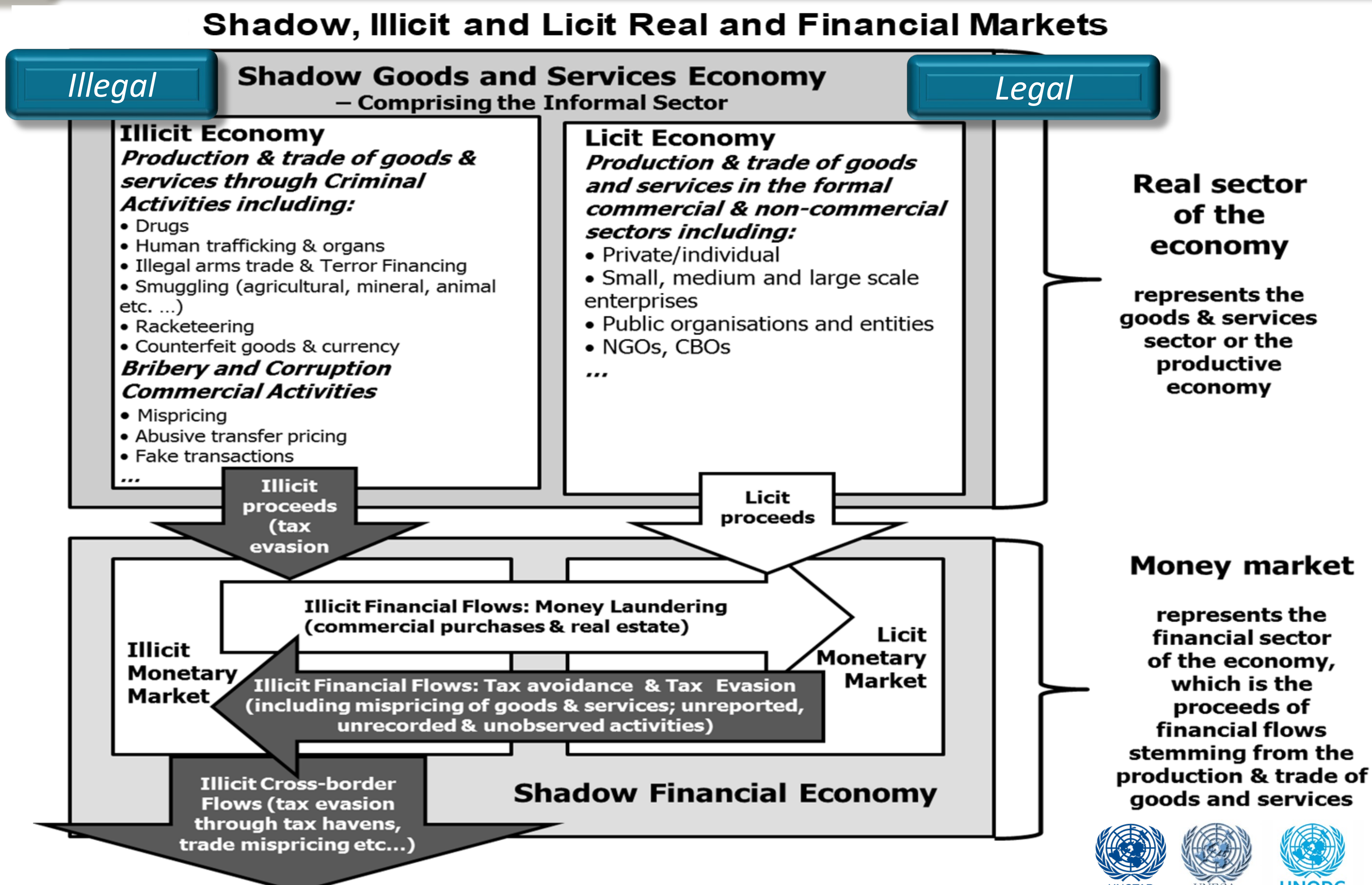
5 Conclusion



UNCTAD UNECA UNODC

Introduction and Background

Measuring a hidden phenomenon



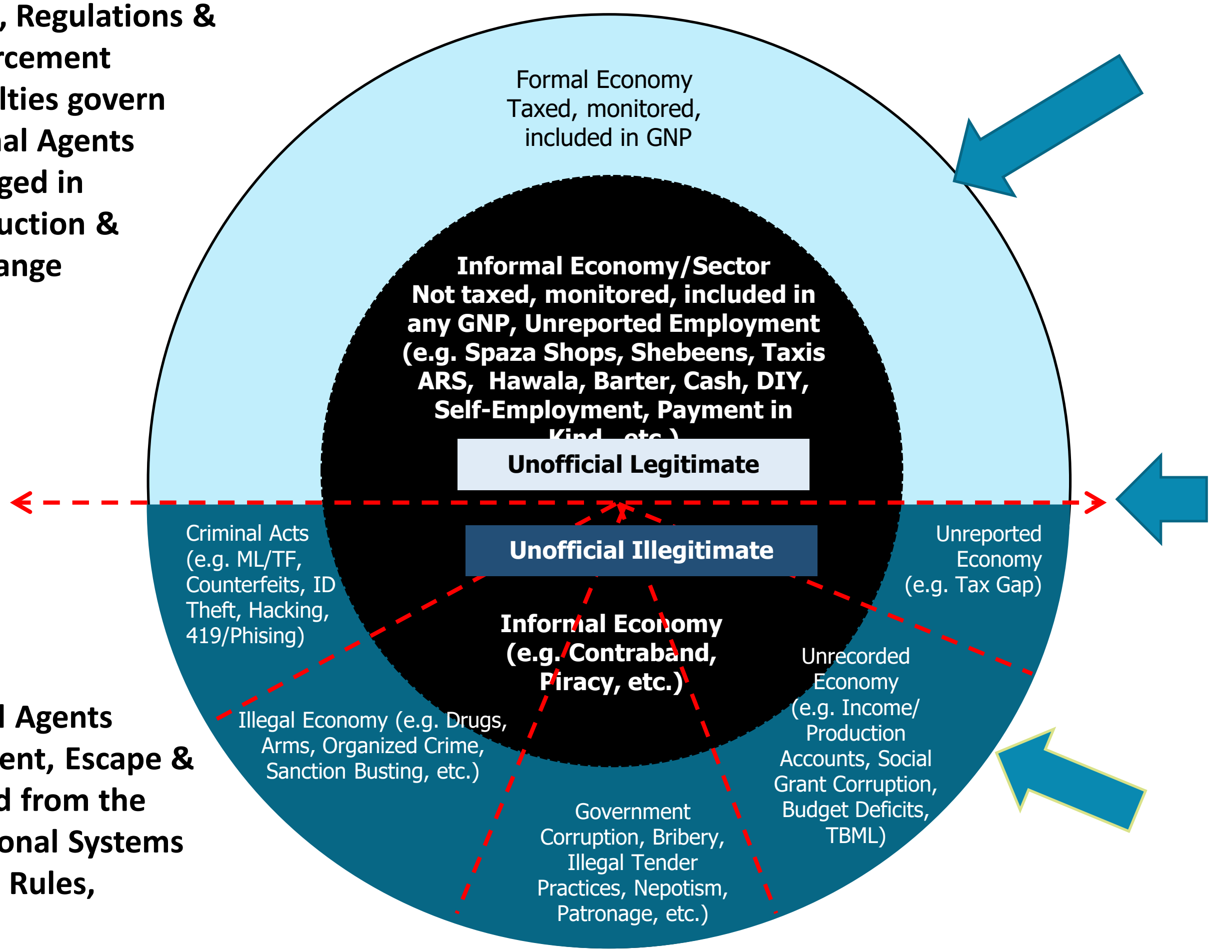
UNCTAD UNECA UNODC

Introduction and Background

Despite considerable research defining IFFs is a challenge

Legal/Licit Regulated

Laws, Regulations & Enforcement Penalties govern Formal Agents engaged in Production & Exchange



Illegal/Illicit Unregulated

Informal Agents Circumvent, Escape & Excluded from the Institutional Systems of Laws, Rules, Rights, Regulations & Enforcement Penalties

__White Market/Economy__:
Formal
Authorized
Transparent
Observed
Reported
Recorded
First

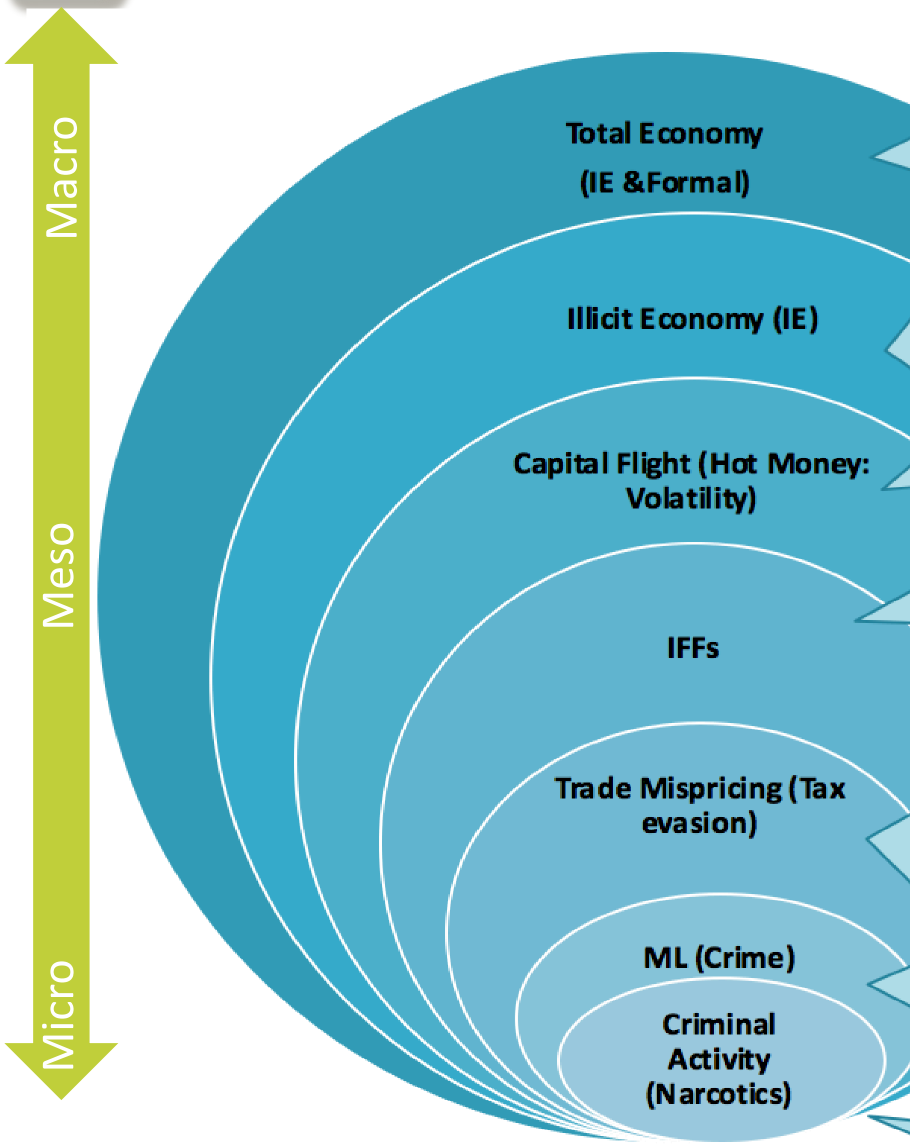
__Grey Market__:
Unauthorized
Shadow
Informal
Unobserved
Unreported
Unrecorded
Second
Opaque

__Black Market__:
Unauthorized
Irregular
Shadow
Informal
Underground
Illegal
Unobserved
Unreported
Unrecorded
Second

Source: South African Financial Intelligence Unit: IFF project, 2012

Introduction and Background

FIC's IFF unit looked at IFFs from the top down and bottom up



Total Economy = IE + Formal Economy

- Conservative CDA estimate of the illegal economy
- MIMIC estimate includes illegal and informal)

Illicit Economy:

- CDA (conservative: Illegal)
- MIMIC (includes illegal and informal). **The WB Columbian Money and Asset Laundering Model estimates the shadow economy*

Hot Money (Capital Flight = Licit + Illicit)

- Capital outflow using the Balance of Payments data

Illicit Financial Flows

- The GFI approach includes **Hot Money plus Trade Mispricing**

Trade Mispricing

- GFI focuses on 35 developed countries plus estimate for ROW trade mispricing
- FIC estimates trade mispricing for all countries due to large role for China. Add Attractiveness weights to explain direction of trade between developing countries (Trade Based Money Laundering)

Money Laundering (Gravity Model) estimates the proceeds of crime and the ML destinations (Threat and risk model). Updated the model with narcotics and trade.
**Columbian Money Asset Laundering Model (WB)*

Criminal Activity: Narcotics, wildlife crime, illicit mining…

The Present

UNECA Paper to Estimating IFFs: Approach

1. Overview

- An overview of the approach describing the high level approach to the problem of quantifying IFFs.

2. Data

- A descrption of the data used in the development of the model and an analysis of how the pros and cons associated with the choice of data.

3. Methodology

- A descrition of the structure and statistical techniques used in the model.

4. Results

- A review of the model results, usually based on the estimated magnitude of IFFs generated by the model.

UNCTAD UNECA UNODC

The Present

The Data Environment

In assessing the data environment for IFFs, four levels of data were identified:

Country Level System of National Accounts Data

Aggregated data based on the System of National Accounts (SNA), which is the internationally agreed standard set of recommendations on how to compile measures of economic activity.

e.g. quarterly basis by the South African Reserve Bank (SARB) with all national accounts information

Country Level Official Statistics Data

Typically aggregated and reported by the statistical bureau of the country, or other entity such as the South African Revenue Authority, representing official statistics for the country. It analyses sectoral data that is supplied to the South African Reserve Bank in order to compile the SNA for the country.

e.g. IMF data, government trade data, UNCOMTRADE data, sectoral production (or statistics on profitability/income after tax) etc.

Commodity/Sector Level (Official) Data

Data collected for specific categories of goods and services or specific industries or from specific government sources. However, this data is analysed and becomes an analytical data set.

e.g. large scale government administrative datasets (or samples thereof); Country-by-Country reports to the OECD; Country Reporting Standard submitted to the OECD; data on gold exports from South Africa or an MNE dataset; Corporate entities tax return datasets; etc.

Transaction Level Data

Data collected at individual transactions level. This data is typically very large and contains significant amounts of detail.

e.g. data collected from banks and other financial institutions, cross border data flows held by the SARB, customs data, tax returns data, etc.

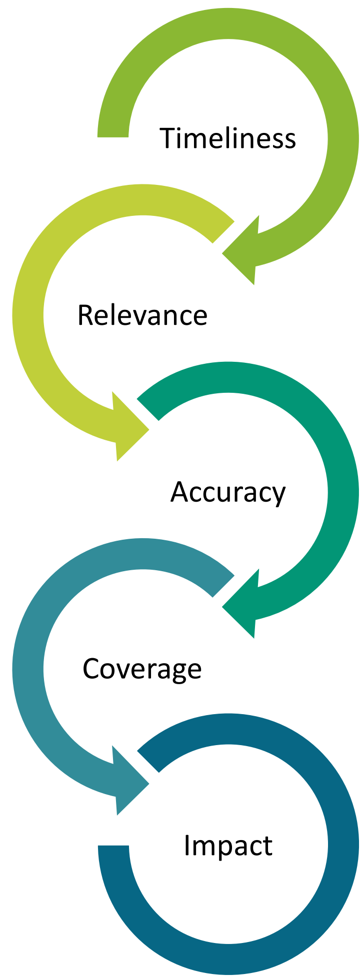
UNCTAD UNECA UNODC

The Present

Evaluating and Scoring the Existing IFF Models

The TRAC-I Scoring Framework

- A scoring framework was developed with the following five core pillars:
    - **T**imeliness;
    - **R**elevance;
    - **A**ccuracy;
    - **C**overage; and
    - **I**mpact
- These five pillars represent the key attributes of a universal IFF indicator. For each pillar, a generic score ranging from 1 (lowest) to 5 (highest) is possible.



- Ease to replicate
- Data availability
- Model structure

- Developing economy context
- Channels of IFF
- Types of IFFs

- Data accuracy
- Statistical Techniques
- Limited assumptions
- Minimal sensitivity

- Multi-country application
- Effort to extend model to other countries

- Generates useful policy info
- Generates powerful diagnostic information

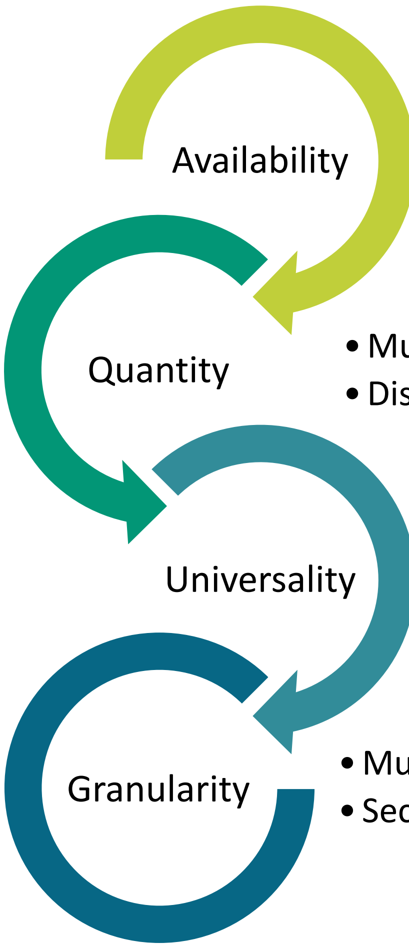
The Present

The IFF Scoring Framework for Data

The key challenge in estimating and measuring IFFs is the availability of compressive, detailed and accurate data.

A Data Quality Framework was also constructed, comprising:

- **Availabilit**y of required data;
- **Quantity** of required data;
- **Universality** of data across countries; and
- **Granularity** of data



- Readily available
- Limited barriers to access

- Multiple years
- Disaggregated

- Available in all countries
- Consistent across countries

- Multiple layers
- Sector/product specific

UNCTAD UNECA UNODC

The Present

Overview of the Finding – Methodology and Data Scoring of IFF Methods

Methodology	Methodology Score							Data Score					OVERALL SCORE (TRAC)	OVERALL SCORE (TRAC-I)
	Timeliness	Relevance	Accuracy	Coverage	Impact	TRAC Score	TRAC-I Score	Availability	Quantity	Universality	Granularity	Data Score		
Trade Estimates														
Country Level Trade Estimates - IMF DOTS														
GFI (Kar (2014), Spanjers and Salomon (2017))	4	3	2	5	3	70%	68%	5	5	5	1	75%	73%	72%
Boyce and Ndikumana (2010)	4	2	2	5	3	65%	64%	4	5	4	1	70%	68%	67%
Bilateral Trade Mis-invoicing (UNCOMTRADE), Nicolaou-Wu (2016)	4	3	3	5	4	75%	76%	4	3	4	3	70%	73%	73%
Trade-Misinvoicing/TBML (UNCOMTRADE) Nicolaou and Wu (2016)	4	3	3	5	4	75%	76%	4	4	4	3	75%	75%	76%
Country Level/Product Level Trade Estimates										3				
Ndikumana (UNCTAD, 2016)(UNCOMTRADE)	4	2	3	5	3	70%	68%	4	3	4	3	70%	70%	69%
Zdanowicz and Pak, Pak and Hong	3	3	4	4	3	70%	68%	4	3	4	3	70%	70%	69%
Bilateral Trade Mis-invoicing (UNCOMTRADE), Nicolaou-Wu (2016)	4	3	3	5	4	75%	76%	4	3	4	3	70%	73%	73%
Trade-Misinvoicing/TBML (UNCOMTRADE) Nicolaou and Wu (2016)	4	3	3	5	4	75%	76%	4	4	4	3	75%	75%	76%
Transaction Level Trade Estimates														
Zdanowicz and Pak, Pak and Hong *same as above)	4	3	4	4	4	75%	76%	4	3	4	4	75%	75%	76%
Shadow Economy - Tax Gap Models														
MIMIC Model	3	3	3	4	3	65%	64%	5	5	5	2	85%	75%	75%
Currency Demand Model (CDA)	5	2	2	5	2	70%	64%	5	5	5	1	80%	75%	72%
SARS Tax Model	5	3	2	5	3	75%	72%	5	5	5	1	80%	78%	76%
Zambian PAYE Tax Gap - Microsimulation	3	4	4	3	3	70%	68%	3	4	3	4	70%	70%	69%
Capital and Wealth														
Hot Money Model	5	2	2	5	2	70%	64%	5	5	5	1	80%	75%	72%
The IMF's Crivelli et al. (2016) model	4	3	2	3	2	60%	56%	3	2	3	1	45%	53%	51%
The UNCTAD (2015) model	4	3	3	3	3	65%	64%	3	3	3	2	55%	60%	60%
The OECD (2015b) model	4	3	4	2	2	65%	60%	2	3	3	3	55%	60%	58%
World Bank Residual Model and Hot Money Model (GFI)	5	3	2	5	3	75%	72%	4	3	5	1	65%	70%	69%
The Clausing (2016) model	3	4	4	2	4	65%	68%	2	4	2	3	55%	60%	62%
The IMF (2014) and EPRS (2015) models	4	3	2	3	2	60%	56%	4	4	3	2	65%	63%	61%
International Tax Avoidance														
The Cobham and Janský (2017) model	4	4	4	4	4	80%	80%	4	4	3	3	70%	75%	75%
The Tørsløv, Wier and Zucman (2017) model	5	4	4	3	4	80%	80%	4	4	5	4	85%	83%	83%
Risk-based Models														
Cobham Risk-based Approach	4	4	4	5	4	85%	84%	4	5	4	2	75%	80%	80%
IFF CRP Risk Tool for African Minstries of Finance	2	5	4	3	5	56%	63%	3	2	3	1	45%	51%	54%
Risk Based Approach – BO-NRA for South Africa	2	4	4	2	5	48%	57%	2	2	3	1	40%	44%	48%
Integrated (IFF) Models														
GFI IFF Estimate (Hot Money Model and Trade Mis-invoicing)	5	3	2	5	3	70%	66%	5	5	5	1	80%	75%	73%
A Composite Indicator to Measure IFFs - SA FIC	4	5	4	5	4	88%	86%	5	4	4	2	74%	81%	80%
Cobham Risk-based Approach (same as above)	4	4	4	5	4	85%	84%	4	5	4	2	75%	80%	80%
Walker Gravity Model (no trade)	3	2	2	5	2	60%	56%	4	3	4	2	65%	63%	61%
Walker Gravity Model (with trade)	3	3	2	5	3	65%	64%	4	3	4	2	65%	65%	65%
The Tørsløv, Wier and Zucman (2017) model	5	4	4	3	4	80%	80%	4	5	5	3	85%	83%	83%
Artificial Intelligence and Machine Learning Models														
Artificial Intelligence and Machine Learning Models	4	4	5	3	5	80%	84%	3	5	4	5	85%	83%	85%

The Future

Introducing Machine Learning and Artificial Intelligence

"Machine learning (ML) is a method of data analysis that automates analytical model building. It is a branch of Artificial Intelligence (AI) based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention - SAS"

- Machine Learning has introduced **Deep Learning** increasing the sophistication of ML and its applications.
- **Financial and Banking sectors use machine learning** to ***fight fraud and money laundering.***
- Research suggests that the ***adoption of machine learning based intelligence*** could drastically ***improve a countries ability to identify, measure and prevent illicit financial flows***.
- Existing machine learning solutions in the ***tax evasion and money laundering space have been developed and implemented***, however this is mainly in ***private sector institutions***.

The Future

ML and AI: A Solution for Estimating and Curbing IFFs in the Future

Traditional Models	Machine Learning (and AI) Models
• Rely on **proven theory to detect known patterns** or **distributions in data**	• Uses **computing power** to **probe the data** for **new and potentially unknown patterns in data**
• **Data modelling** has been restricted to the **detection of theoretically known phenomena** within a dataset. The aim was **to fit a model to the data.**	• With AI technology and Machine Learning, it is possible to **sift through massive amounts of data in search of the unknown.**
• **Illicit transactions** are designed to **pass undetected as far as possible**, making **their detection difficult, time consuming and case-specific**	• Machine Learning provides a potential solution to this challenge where **known and unknown patterns can be detected**, including the **slightest deviation from these patterns**, and can do so at **speeds unattainable by human beings or even traditional statistical models.**
	• By training machine learning models to **recognise known patterns and unknown patterns** of illicit trade, there is **huge potential for improved IFF detection, quantification and eventual reduction**. • ML techniques can augment/upgrade existing IFF models

...ture

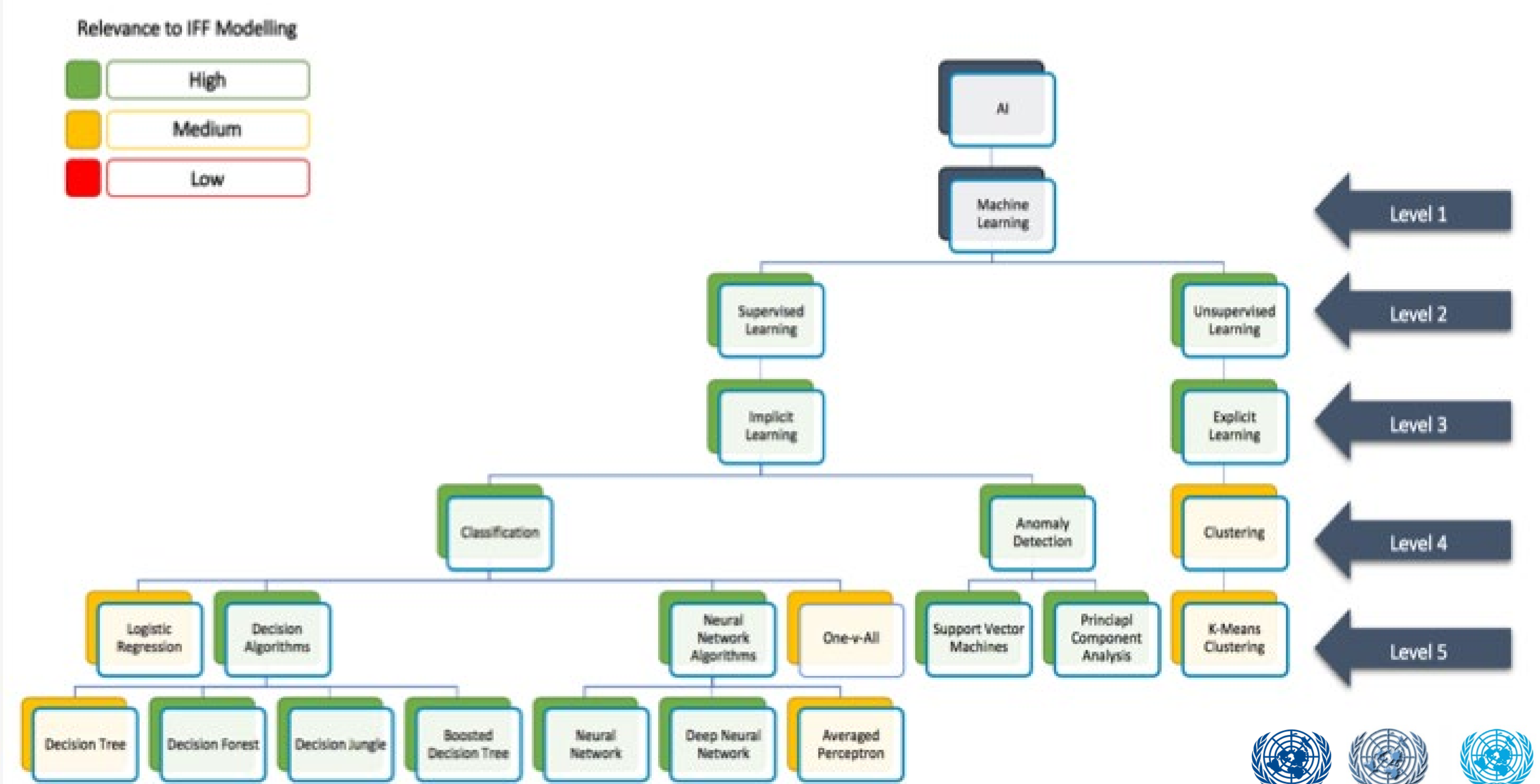
Machine Learning – Understanding the Elements and Appropriateness for IFFs

Level 1	**Which branches of AI are most suitable for IFF modelling?** (Machine Learning)
Level 2	**Which approach to machine learning is most suitable for IFF modelling?**
Level 3	**Which types of learning are most suitable for IFF modelling?**
Level 4	**Which applications of machine learning are most suitable for IFF modelling?**
Level 5	**Which machine learning algorithms are most suitable for IFF modelling?**

The Future

Machine Learning – Understanding the Elements

- *The core feature of any Machine Learning model is the **ability to learn**.*
- *This learning can take place **explicitly, where a body of knowledge is taught to the machine, or implicitly, where the machine begins to learn independently of the explicit information**. This forms the **primary difference between supervised and unsupervised learning. Learning takes place at different stages of a models lifecycle,***



UNCTAD UNECA UNODC

The Future

Examples of Machine Learning Solutions in the IFF Space

There are a few examples of ML solutions applied to IFFs as shown below. Bare in mind that these solutions are geared towards detection of IFFs and therefore operate in real time, detecting illicit transactions as they occur. These solutions can be used to identify illicit transactions as they occur or can be applied to historical data to measure the magnitude of IFFs

Case Study 1

- **Title:** Credit card fraud forecasting model based on clustering analysis and integrated support vector machine
- **Author:** Chunhua Wang and Dong Han
- **Year:** 2018
- **Country:** China
- **Publisher:** School of Animation, Huanghuai University, Zhumadian, Henan

The results of the study indicate that the use of **machine learning algorithms are suitable for the detection and forecasting of credit card fraud.**

Case Study 2

- **Title:** Neural Network Analysis of International Trade
- **Author:** Isaac Wohl and Jim Kennedy
- **Year:** 2018
- **Country:** United States
- **Publisher:** U.S. International Trade Commission (USITC)

It is expected that neural networks would have **equal or better predictive accuracy** than the other estimators, since neural networks with the right specifications can approximate any continuous function, linear or non-linear.

Case Study 3

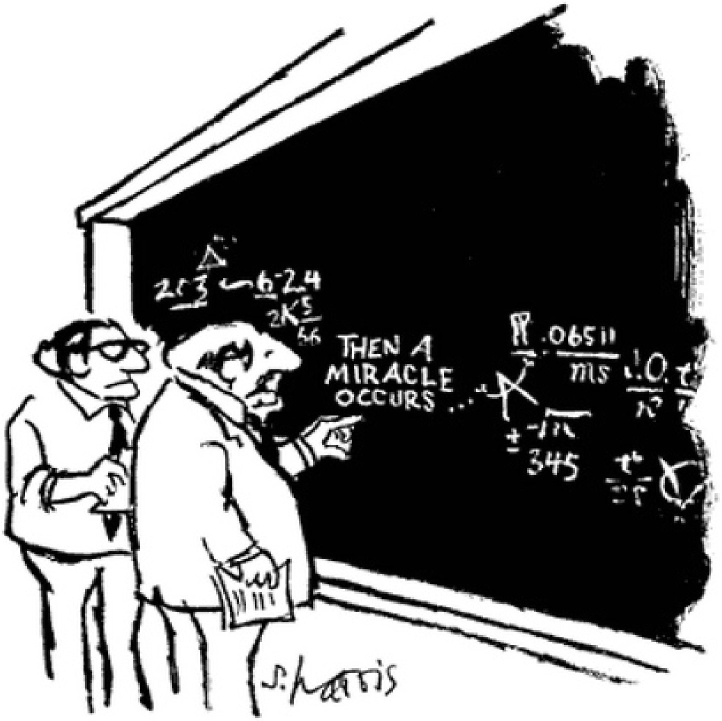
- **Title:** Neural network models: Foundations and applications to an audit decision problem
- **Author:** Rebecca C. Wu
- **Year:** 1997
- **Country:** Republic of China
- **Publisher:** National Taiwan University, Taipei, Taiwan

The results indicated that **the neural network was able to predict potential tax abuse cases with 94 percent accuracy using a two-layered model and with 95 percent accuracy using a three-layered model.**

The Roadmap to Indicator 16.4.1

Lessons for Developing Indicator 16.1.4 from Heading the IFF Unit at SA's FIC

Does this feel familiar?



"I think you should be more explicit here in step two."

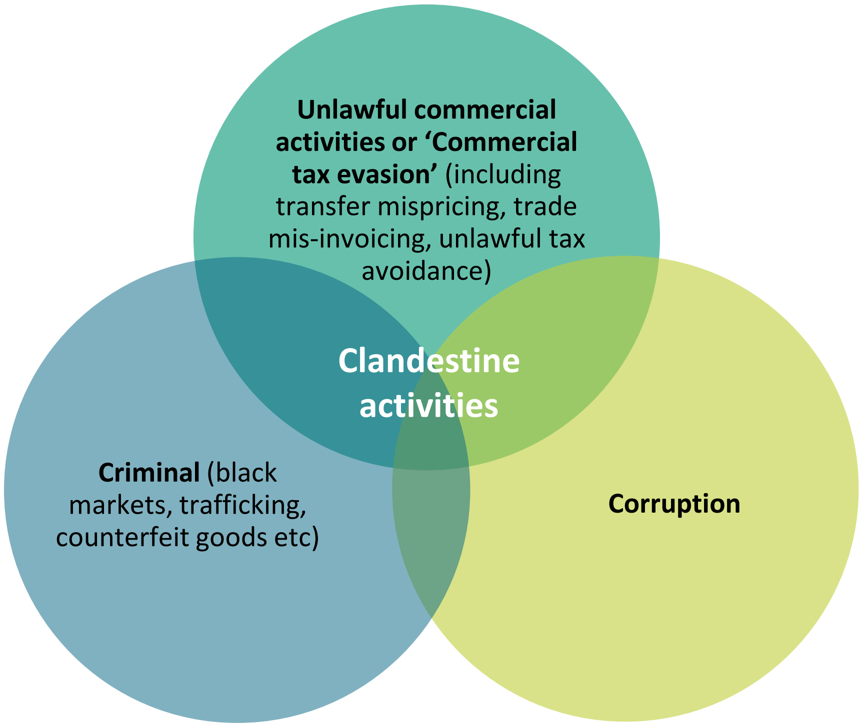
- **Lesson 1:** Definition is key but don't get stuck there
- **Lesson 2: Multiplicity** – we used multiple models, data sources, multiple approaches, levels of modelling, etc. ***Think composite/multi-level latent variable indicator***
- **Lesson 3:** Look at the productive and monetary sides of the economy when measuring: This might be one of the most important factors, especially for trade, tax evasion, criminal economy, etc.
- **Lesson 4:** Think top-down and bottom up.
- **Lesson 5:** Some measure is better than no measure…
- **Lesson 6:** We weren't allowed to use transactional/administrative data, despite setting up MOUs; that didn't stop us. Use all data sources available and proxies and alternate data sources…
- **Lesson 7:** You will always be criticized … it's a latent variable
- **Lesson 8:** Pilot, test, iterate, adapt and repeat…

UNCTAD UNECA UNODC

The Roadmap to Indicator 16.4.1

Recommendations in Developing Indicator 16.4.1

Primary Challenge in Estimating IFFs:



- Guided by your Theory of Change, empower governments to **proactively curb IFFs**
- **Officially published data is king**
- **Estimates are relevant in developing countries where transactional, administrative data is not available**.
- **Estimates for the future? Pilot of an Artificial Intelligence and Machine Learning** model using granular, transactional and administrative data with Deep Learning.
- **Go for a composite indicator** but take care not to double count. Use FATF, other indicators or risk models
- Good methodologies with limited or analytical databases ought to be avoided
- Tørsløv et al is *a good models using SNA Macro data.*

- **Risk models** are useful from a policy prescript/diagnostic perspective.
- The MIMIC, Hot Money and Currency Demand models are estimates but perform well because their **simplicity, access to globally accessible time series data and the ease of replicability.**
- The trade mis-invoicing models (transactional) by Zdanowicz and Pak perform well. Nicolaou and Wu models **address double counting** and provide **more useful policy insights and diagnostic capabilities.**

UNCTAD UNECA UNODC

The Journey - Roadmap

Implementing an AI and ML Pilot

1. Data Quality Assurance

- IMF vs UNCTAD datasets
- Variances of results due to different reporting standards
- Implications of poor data quality for ML (garbage in – garbage out)

2. Data Granularity

- Progression from high level data (BoP, National Accounts etc.) to transaction level data (customs, SARS, Banks etc.)
- The greater the volume of data, the better (for ML). This is assuming that data quality is acceptable
- Speak to live anomaly detection to identify and prevent IFFs.
  This requires a robust relational dataset between the various entities in a country (customs, SARS, banks etc.)

3. Identifying Proxies

- a) generate data where it is missing
  In most cases, data availability is an issue. To address this issue it is useful to use proxies to estimate missing data.
- b) validate data and estimates.
  Where data has been aggregated or estimated, proxies are useful to test results against other data at a high level
- c) better understand data
  Even if a data source is trusted and validated, the use of external datasets to better understand or further disaggregate the data is often useful

4. Implement Live Pilots that draw on real time/live data sources

Conclusions on the Journey/Road Map

Roadmap from Today to Measuring IFFs in 2030

- Most of the Existing Models are Problematic because of the **Underlying Data**
- In the absence of Machine Learning Models, the **existing models will do**
- A KEY to success is to PILOT Machine Learning Models in developing countries using transactional data … this will change the indicator in the future.
- My take as a policy practitioner: Develop a latent **multi-level or composite indicator with discrete elements** to include:

1. **Measure: Trade mis-invoicing measure** (Nicolaou-Wu TBML model or other improved versions relying on Comtrade) (Country Level official data) OR a transactional model (Zdanowicz&Pak
2. **Measure: Add GFI's World Bank Residual Component** (for a Balance of Payment's analysis)
3. **Measure: Add a tax evasion/tax shifting** component (Cobham and Jansky, Torslov etc.)
4. **Risk Index/Model:** Add a risk indicator either from: Cobham's Risk Model OR GIZ's CRA Tool
5. **Measure: Add a criminal economy measure** (UNODC work or Walker's Gravity Model or Currency Demand Model (latent estimate of black money)
6. **Index: Add a risk/transparency index for corruption**
7. Consider a composite index: FATF's Mutual Evaluations or 40+ recommendations, FSI, Walker's Trade Gravity Model and ML Attractiveness Index

UNCTAD UNECA UNODC