Benford testing procedure validation

Andrea Cerasa

JRC I.3 – Text and Data Mining Unit
Statistics and Information Technologies for Anti-Fraud, Security and Trade

Benford's law conference, Stresa, 10-12 July 2019

Objective: study and assess the properties of the Benford testing procedures described in

"Newcomb–Benford law and the detection of frauds in international trade" *Cerioli, A., Barabesi, L., Cerasa, A., Menegatti, M., & Perrotta, D.* (2019). *Proceedings of the National Academy of Sciences, 116(1), 106-115.*

in terms of ***size***, ***power*** and ***rate of false signals*** on a real environment.

Objective: study and assess the properties of the Benford testing procedures described in

"Newcomb–Benford law and the detection of frauds in international trade" *Cerioli, A., Barabesi, L., Cerasa, A., Menegatti, M., & Perrotta, D.* (2019). *Proceedings of the National Academy of Sciences, 116(1), 106-115.*

in terms of ***size***, ***power*** and ***rate of false signals*** on a real environment.

In the article, the proposed procedure is shown to work properly in simulated markets.

Objective: study and assess the properties of the Benford testing procedures described in

"Newcomb–Benford law and the detection of frauds in international trade" *Cerioli, A., Barabesi, L., Cerasa, A., Menegatti, M., & Perrotta, D.* (2019). *Proceedings of the National Academy of Sciences, 116(1), 106-115.*

in terms of ***size***, ***power*** and ***rate of false signals*** on a real environment.

In the article, the proposed procedure is shown to work properly in simulated markets.

However, no data were available for a huge validation experiment.

Description of the proposed procedure:

Trader X:
N transactions
on
M distinct products

- N_1 *transactions on Product 1*
- N_2 *transactions on Product 2*
- …
- N_M *transactions on Product M*

Description of the proposed procedure:

Trader X:
N transactions
on
M distinct products

- N_1 *transactions on Product 1*
- N_2 *transactions on Product 2*
- …
- N_M *transactions on Product M*

χ^2 on the first digit of the N statistical values

Description of the proposed procedure:

Trader X:
N transactions
on
M distinct products

- N_1 *transactions on Product 1*
- N_2 *transactions on Product 2*
- …
- N_M *transactions on Product M*

χ^2 on the first digit of the N statistical values

Pseudo-Trader X:
N pseudo-transactions
on
M distinct products

- N_1 *pseudo-transactions on Product 1*
- N_2 *pseudo-transactions on Product 2*
- …
- N_M *pseudo-transactions on Product M*

Description of the proposed procedure:

Trader X:
N transactions
on
M distinct products

- N_1 *transactions on Product 1*
- N_2 *transactions on Product 2*
- …
- N_M *transactions on Product M*

χ^2 on the first digit of the N statistical values

Pseudo-Trader X:
N pseudo-transactions
on
M distinct products

- N_1 *pseudo-transactions on Product 1*
- N_2 *pseudo-transactions on Product 2*
- …
- N_M *pseudo-transactions on Product M*

10,000 times (or more)

Description of the proposed procedure:

Trader X:
N transactions
on
M distinct products

- N_1 *transactions on Product 1*
- N_2 *transactions on Product 2*
- …
- N_M *transactions on Product M*

χ^2 on the first digit of the N statistical values

Pseudo-Trader X:
N pseudo-transactions
on
M distinct products

- N_1 *pseudo-transactions on Product 1*
- N_2 *pseudo-transactions on Product 2*
- …
- N_M *pseudo-transactions on Product M*

10,000 times (or more)

Empirical distribution of the χ^2 on the first digit for **Trader X**

Description of the proposed procedure:

Trader X:
N transactions
on
M distinct products

- N_1 *transactions on Product 1*
- N_2 *transactions on Product 2*
- …
- N_M *transactions on Product M*

Pseudo-Trader X:
N pseudo-transactions
on
M distinct products

- N_1 *pseudo-transactions on Product 1*
- N_2 *pseudo-transactions on Product 2*
- …
- N_M *pseudo-transactions on Product M*

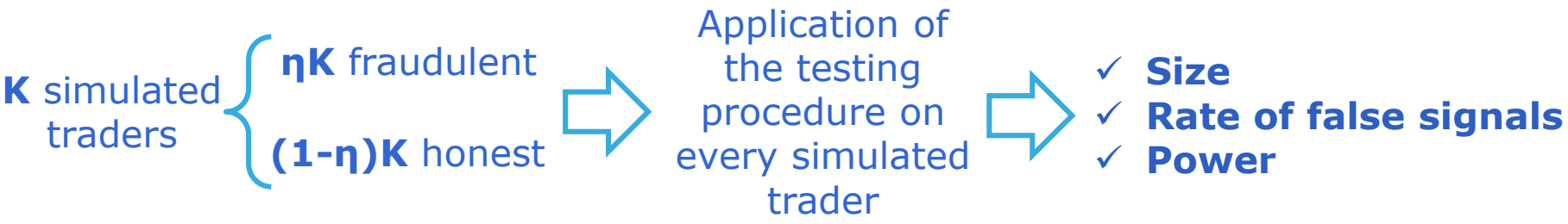
10,000 times (or more)

χ^2 on the first digit of the N statistical values

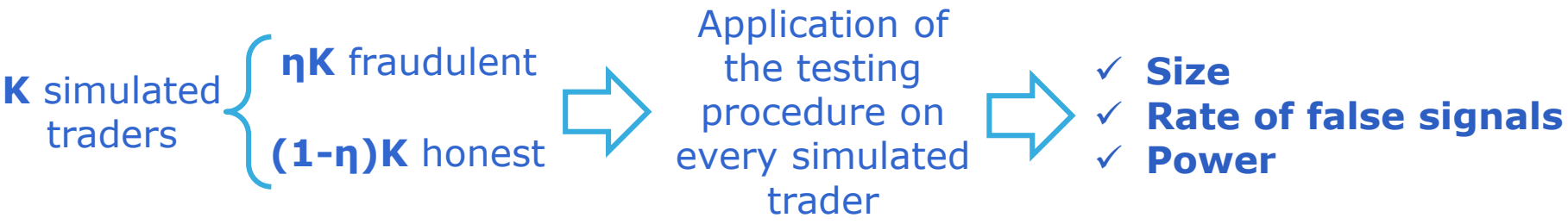
Empirical distribution of the χ^2 on the first digit for **Trader X**

Exact p-value

Simulated markets were used to assess the properties of the procedure:

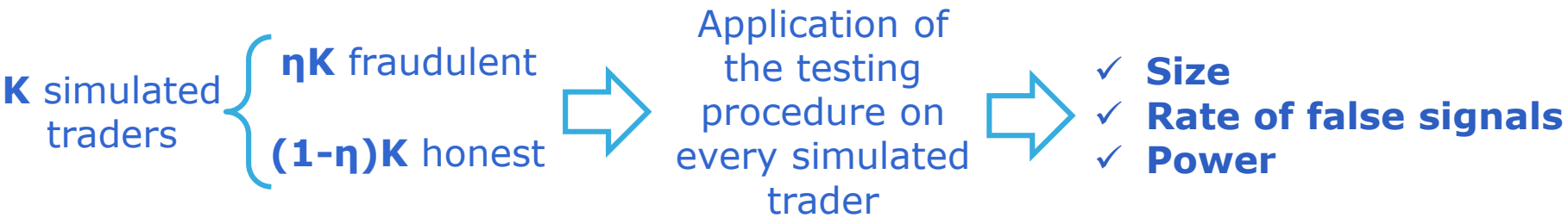


Simulated markets were used to assess the properties of the procedure:



For a complete view of the performances of the test, in the simulation experiments we consider:

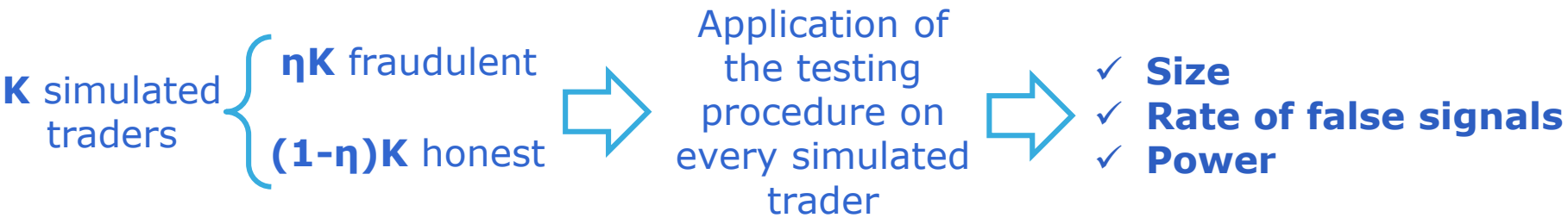
Simulated markets were used to assess the properties of the procedure:



For a complete view of the performances of the test, in the simulation experiments we consider:

- ✓ Different combinations of **N** (number of transactions) and **M** (number of distinct products);

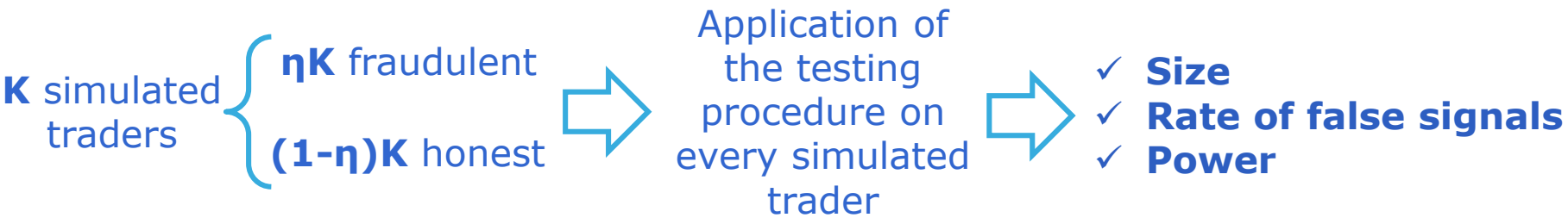
Simulated markets were used to assess the properties of the procedure:



For a complete view of the performances of the test, in the simulation experiments we consider:

- ✓ Different combinations of **N** (number of transactions) and **M** (number of distinct products);
- ✓ Different fraud strategies and fraud intensities *(defined as the percentage of manipulated transactions)*;

Simulated markets were used to assess the properties of the procedure:



For a complete view of the performances of the test, in the simulation experiments we consider:

- ✓ Different combinations of **N** (number of transactions) and **M** (number of distinct products);
- ✓ Different fraud strategies and fraud intensities *(defined as the percentage of manipulated transactions)*;
- ✓ Different fractions of fraudulent traders **η**.

Collaboration with Austrian Customs authorities that provided us the necessary data for the back-testing experiment in a real context.

The dataset provided included all the trades involving products sensible to under-evaluation strategies in the period between 2011 and 2018.

Collaboration with Austrian Customs authorities that provided us the necessary data for the back-testing experiment in a real context.

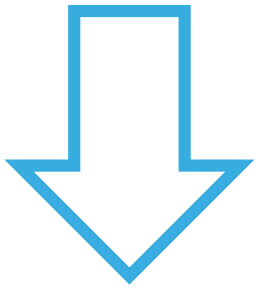
The dataset provided included all the trades involving products sensible to under-evaluation strategies in the period between 2011 and 2018.

9,729,946 records
(399,445 importers)

Collaboration with Austrian Customs authorities that provided us the necessary data for the back-testing experiment in a real context.

The dataset provided included all the trades involving products sensible to under-evaluation strategies in the period between 2011 and 2018.

9,729,946 records
(399,445 importers)

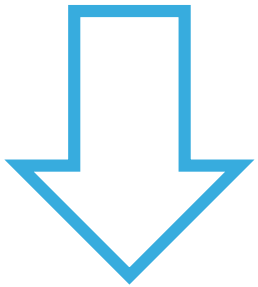


55,710 checked records
(9,570 importers)

Collaboration with Austrian Customs authorities that provided us the necessary data for the back-testing experiment in a real context.

The dataset provided included all the trades involving products sensible to under-evaluation strategies in the period between 2011 and 2018.

9,729,946 records
(399,445 importers)



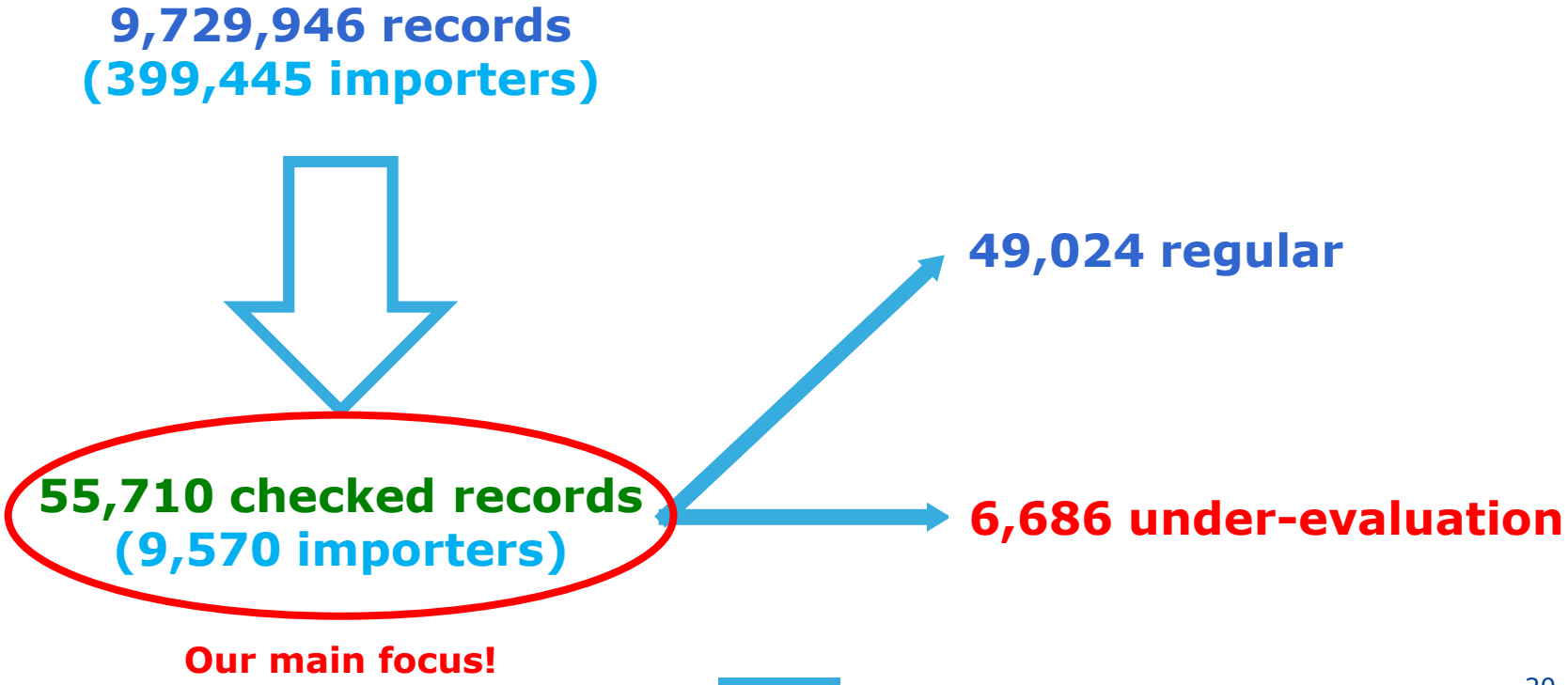
55,710 checked records
(9,570 importers)

49,024 regular

6,686 under-evaluation

Collaboration with Austrian Customs authorities that provided us the necessary data for the back-testing experiment in a real context.

The dataset provided included all the trades involving products sensible to under-evaluation strategies in the period between 2011 and 2018.



For a reliable application of our testing procedure, we need:

For a reliable application of our testing procedure, we need:

1) Traders with at least 50 transactions

55,710 checked records
(9,570 importers)

39,245 checked records
(3,856 importers)

For a reliable application of our testing procedure, we need:

1) Traders with at least 50 transactions

55,710 checked records
(9,570 importers)

39,245 checked records
(3,856 importers)

2) A clear distinction between <u>honest</u> and <u>fraudulent</u> traders

Not easy!

For a reliable application of our testing procedure, we need:

1) Traders with at least 50 transactions

55,710 checked records (9,570 importers) → 39,245 checked records (3,856 importers)

2) A clear distinction between <u>honest</u> and <u>fraudulent</u> traders

Not easy!

- AT checks focus on single transactions
- Our method points at traders

For a reliable application of our testing procedure, we need:

1) Traders with at least 50 transactions

55,710 checked records (9,570 importers) → 39,245 checked records (3,856 importers)

2) A clear distinction between <u>honest</u> and <u>fraudulent</u> traders

Not easy!

- AT checks focus on single transactions
- Our method points at traders

It is difficult to give a reliable definition of an honest trader

Transactions considered	all	9,729,946
Checked companies considered	traders with at least 50 transactions & at least 1 transaction checked	3,856
Definition of honest trader	no fraud detected	3,337
Definition of fraudulent trader	at least 1 fraud detected	519

Transactions considered	all	9,729,946
Checked companies considered	traders with at least 50 transactions & at least 1 transaction checked	3,856
Definition of honest trader	no fraud detected	3,337
Definition of fraudulent trader	at least 1 fraud detected	519

Trader A

100 transactions

99 not-checked

1 checked

Regular

Transactions considered	all	9,729,946
Checked companies considered	traders with at least 50 transactions & at least 1 transaction checked	3,856
Definition of honest trader	no fraud detected	3,337
Definition of fraudulent trader	at least 1 fraud detected	519

Trader A	**Trader B**
100 transactions	100 transactions
99 not-checked	35 not-checked
1 checked	65 checked
Regular	*All regular*

Transactions considered	all	9,729,946
Checked companies considered	traders with at least 50 transactions & at least 1 transaction checked	3,856
Definition of honest trader	no fraud detected	3,337
Definition of fraudulent trader	at least 1 fraud detected	519

Trader A	**Trader B**	**Trader C**
100 transactions	100 transactions	100 transactions
99 not-checked 1 checked	35 not-checked 65 checked	99 not-checked 1 checked
Regular	*All regular*	*1 under-eval.*

Transactions considered	all	9,729,946
Checked companies considered	traders with at least 50 transactions & at least 1 transaction checked	3,856
Definition of honest trader	no fraud detected	3,337
Definition of fraudulent trader	at least 1 fraud detected	519

Trader A	**Trader B**	**Trader C**	**Trader D**
100 transactions	100 transactions	100 transactions	100 transactions
99 not-checked 1 checked	35 not-checked 65 checked	99 not-checked 1 checked	35 not-checked 65 checked
Regular	*All regular*	*1 under-eval.*	*50 under-eval.*

Transactions considered	all	9,729,946
Checked companies considered	traders with at least 50 transactions & at least 1 transaction checked	3,856
Definition of honest trader	no fraud detected	3,337
Definition of fraudulent trader	at least 1 fraud detected	519

Trader A	**Trader B**	**Trader C**	**Trader D**
100 transactions	100 transactions	100 transactions	100 transactions
99 not-checked 1 checked	35 not-checked 65 checked	99 not-checked 1 checked	35 not-checked 65 checked
Regular	*All regular*	*1 under-eval.*	*50 under-eval.*

Results will then consider different definitions of the honest traders set.

Filter 1

Improve the reliability of honest traders definition by considering only the traders with a significant fraction (40%) of transactions checked.

Filter 1

Improve the reliability of honest traders definition by considering only the traders with a significant fraction (40%) of transactions checked.

Transactions considered	all	9,729,946
Checked companies considered	traders with at least 50 transactions & at least **40%** of transactions checked	37
Definition of honest trader	no fraud detected	25
Definition of fraudulent trader	at least 1 fraud detected	12

Filter 1

Improve the reliability of honest traders definition by considering only the traders with a significant fraction (40%) of transactions checked.

Transactions considered	all	9,729,946
Checked companies considered	traders with at least 50 transactions & at least **40%** of transactions checked	37
Definition of honest trader	no fraud detected	25
Definition of fraudulent trader	at least 1 fraud detected	12

Size: 8.00%
Rate of False Signals: 18.18%
Power: 75.00%

Filter 1

Improve the reliability of honest traders definition by considering only the traders with a significant fraction (40%) of transactions checked.

Transactions considered	all	9,729,946
Checked companies considered	traders with at least 50 transactions & at least **40%** of transactions checked	37
Definition of honest trader	no fraud detected	25
Definition of fraudulent trader	at least 1 fraud detected	12

Size: 8.00%
Rate of False Signals: 18.18%
Power: 75.00%

Nice results but too few traders…

Filter 2

Focus only on the 55,710 checked records. In this way, it is possible to exactly discriminate honest and fraudulent traders, since all transactions have been checked.

Filter 2

Focus only on the 55,710 checked records. In this way, it is possible to exactly discriminate honest and fraudulent traders, since all transactions have been checked.

Transactions considered	Only checked	55,710
Checked companies considered	traders with at least 50 transactions	143
Definition of honest trader	no fraud detected	82
Definition of fraudulent trader	at least 1 fraud detected	61

Filter 2

Focus only on the 55,710 checked records. In this way, it is possible to exactly discriminate honest and fraudulent traders, since all transactions have been checked.

Transactions considered	Only checked	55,710
Checked companies considered	traders with at least 50 transactions	143
Definition of honest trader	no fraud detected	82
Definition of fraudulent trader	at least 1 fraud detected	61

Size: 7.32%
Rate of False Signals: 31.58%
Power: 21.31%

Filter 2

Focus only on the 55,710 checked records. In this way, it is possible to exactly discriminate honest and fraudulent traders, since all transactions have been checked.

Transactions considered	Only checked	55,710
Checked companies considered	traders with at least 50 transactions	143
Definition of honest trader	no fraud detected	82
Definition of fraudulent trader	at least 1 fraud detected	61

Size: 7.32%
Rate of False Signals: 31.58%
Power: 21.31%

Good results in general.
Low power, but also low fraud intensity (median: 12.5%)
Still high size and high rate of false signals…
Further investigations are required…

Analyzing in details the statistical values declared by the 143 traders, we observed that:

Analyzing in details the statistical values declared by the 143 traders, we observed that:

✓ Some of them declared values in a very small range;

$$Range = Log_{10}\big(Max(V)\big) - Log_{10}(Min(V))$$

Analyzing in details the statistical values declared by the 143 traders, we observed that:

- ✓ Some of them declared values in a very small range;

$$Range = Log_{10}\big(Max(V)\big) - Log_{10}(Min(V))$$

- ✓ Some of them repeated the same purchase structure (same product, same origin, same quantity, same value,…) a significant number of times.

Analyzing in details the statistical values declared by the 143 traders, we observed that:

- ✓ Some of them declared values in a very small range;

$$Range = Log_{10}\big(Max(V)\big) - Log_{10}(Min(V))$$

- ✓ Some of them repeated the same purchase structure (same product, same origin, same quantity, same value,…) a significant number of times.

Both these features may affect the results obtained for the proposed procedure, especially in term of size.

Analyzing in details the statistical values declared by the 143 traders, we observed that:

- ✓ Some of them declared values in a very small range;

$$Range = Log_{10}\big(Max(V)\big) - Log_{10}(Min(V))$$

- ✓ Some of them repeated the same purchase structure (same product, same origin, same quantity, same value,…) a significant number of times.

Both these features may affect the results obtained for the proposed procedure, especially in term of size.

Therefore we decided to further filter the traders…

Filter 3

In particular, among the 143 traders of the previous filter, we focus only on the ones that:

Filter 3

In particular, among the 143 traders of the previous filter, we focus only on the ones that:

Transactions considered	Only checked	55,710
Checked companies considered	traders with at least 50 transactions & Declared Statistical Value Range>3 & Percentage of Repeated Purchases<20%	102
Definition of honest trader	no fraud detected	63
Definition of fraudulent trader	at least 1 fraud detected	39

Filter 3

In particular, among the 143 traders of the previous filter, we focus only on the ones that:

Transactions considered	Only checked	55,710
Checked companies considered	traders with at least 50 transactions & Declared Statistical Value Range>3 & Percentage of Repeated Purchases<20%	102
Definition of honest trader	no fraud detected	63
Definition of fraudulent trader	at least 1 fraud detected	39

Size: 3.17%
Rate of False Signals: 25.00%
Power: 15.38%

Filter 3

In particular, among the 143 traders of the previous filter, we focus only on the ones that:

Transactions considered	Only checked	55,710
Checked companies considered	traders with at least 50 transactions & Declared Statistical Value Range>3 & Percentage of Repeated Purchases<20%	102
Definition of honest trader	no fraud detected	63
Definition of fraudulent trader	at least 1 fraud detected	39

Size: 3.17%
Rate of False Signals: 25.00%
Power: 15.38%

Size very close to the nominal level (1%)
Improved rate of false signals
Lower power, but also lower fraud intensity (median: 3.9%)

Concluding remarks

Concluding remarks

- The different target of the checks (single transaction vs trader) affects the definition of honest traders and significantly reduces the dimension of the validation set

Concluding remarks

- The different target of the checks (single transaction vs trader) affects the definition of honest traders and significantly reduces the dimension of the validation set

- Power seems poor but is directly proportional to the intensity of fraud, which is usually low.

Concluding remarks

- The different target of the checks (single transaction vs trader) affects the definition of honest traders and significantly reduces the dimension of the validation set
- Power seems poor but is directly proportional to the intensity of fraud, which is usually low.
- When a suitable filter is applied, the size of the testing procedure is close to its nominal level.

Concluding remarks

- The different target of the checks (single transaction vs trader) affects the definition of honest traders and significantly reduces the dimension of the validation set
- Power seems poor but is directly proportional to the intensity of fraud, which is usually low.
- When a suitable filter is applied, the size of the testing procedure is close to its nominal level.
- An heuristic filter (i.e. filter 3) guarantees the best results in this sense.

Concluding remarks

- The different target of the checks (single transaction vs trader) affects the definition of honest traders and significantly reduces the dimension of the validation set
- Power seems poor but is directly proportional to the intensity of fraud, which is usually low.
- When a suitable filter is applied, the size of the testing procedure is close to its nominal level.
- An heuristic filter (i.e. filter 3) guarantees the best results in this sense.
- A refined simulation scheme taking into account similarities and dependences of traders behavior will be object of future studies