Social Inspections based on Risk Profiles

Dacian Sirbu, Agency for Social Benefits and Inspection, Romania
Vlad Grigoras, Extended Term Consultant, World Bank

Session 7, Romania

WORKSHOP ON
REDUCING ERROR, FRAUD & CORRUPTION (EFC) IN SOCIAL PROTECTION PROGRAMS
June 8-12, 2014, Opatija, Croatia

Outline

- **Why** inspections based on risk-profiling
- **How** to do inspections based on risk-profiling
- **Practical aspects** of doing inspections based on risk-profiling
- **Plans** for risk-profiling in Romania

Untargeted social inspections…

Assuming that:

- Error fraud and corruption (EFC) occurs in roughly 5% of cases and closer to 10% when benefits are more complex (e.g. means-tested)
- The average benefit in Romania is USD100
- We inspect 1,000 beneficiaries (cost per inspection – USD10)

Cost	USD10*1000 inspections	USD 10,000
Benefit	USD100*50 cases of fraud	USD 5,000
Cost-benefit ratio		10,000: 5,000 = 2: 1

- The benefits are lower than the costs
- 95% of beneficiaries are compliant, and will be inconvenienced by the activity of social inspectors

Solution: Move to risk-based investigations

Targeted social inspections…

Assuming that:

- We increase the probability of detecting EFC at 50% of the total number of social inspections

Cost	USD10*1000 inspections	USD 10,000
Benefit	USD100*500 cases of fraud	USD 50,000
Cost-benefit ratio		10,000: 50,000 = 1: 5

We need to find ways to target the social inspection campaigns on the beneficiaries with a higher risk of EFC.

Potential Solutions to target social inspections

Data matching	**Events of noncompliance that have already occurred are identified.** It is problematic when: ❑ There are no other databases to match (lack of protocols) or the databases do not contain information relevant for detecting fraud ❑ Different databases measure different things or there is no unique ID
Fraud referral	**A hot line where people can uncover cases of fraud/error.** ❑A dedicated team needs to be in place. ❑Success depends on local social values/norms. ❑A long term investment in information campaigns is needed. ❑Only some types of fraud/error can be uncovered.
Manual screening	**Manual selection of cases by inspectors based on their own knowledge of beneficiaries' behavior and environment.** ❑It increases the risk of corruption. ❑Social inspectors can miss some aspects of noncompliance.

Alternative solution: Move to risk-based investigations

What is Risk Profiling?

Set of statistical procedures that would allow the social inspectors to better identify the cases with a higher probability of EFC

Risk-scoring techniques comparable to those used to categorize clients in banking or insurance.

Advantages:

- It increases the cost-effectiveness of social inspections
- It reduces the length of time fraud stays in the system
- It decreases the number of inspections for compliant beneficiaries

Outline

- **Why** inspections based on risk-profiling
- **How** to do inspections based on risk-profiling
- **Practical aspects** of doing inspections on risk-profiling
- **Plans** for risk-profiling in Romania

Steps when doing inspections based on risk profiles

1. Build a dataset with the population of beneficiaries and their characteristics
2. Select a random sample of beneficiaries
3. Perform inspections on the random sample of beneficiaries
4. Identify the cases in the sample with higher probability of EFC based on their characteristics.
5. Identify in the total population the cases having the characteristics that flagged a higher risk of EFC
6. Perform inspections primarily on the cases in the population that show higher risk of EFC
7. Review the model based on new iterations

1. Build a dataset with the population of beneficiaries and their characteristics

Population of beneficiaries

No	Area	No of members	etc.
1	Urban	2	
2	Rural	3	
3	Rural	1	
4	Rural	4	
5	Rural	7	
6	Urban	2	
7	Urban	4	
.....			
.....			
100,000	Rural	1	

Types of characteristics: likely to predict fraud/error and available for each beneficiary

Examples:

- Residential area/type of locality
- No of members/no of children
- No of members of active age
- Maximum level of education
- Amount of declared incomes
- Type of family (lone parents etc.)
- Period in the program
- Health status
- Information at local level

2. Select a random sample of beneficiaries

Population of beneficiaries

No	Area	No of members	etc.
1	Urban	2	
2	Rural	3	
3	Rural	1	
4	Rural	4	
5	Rural	7	
6	Urban	2	
7	Urban	4	
…..			
…..			
100,000	Rural	1	

Sample of beneficiaries

No	Area	No of members	etc.
1	Rural	1	
2	Rural	4	
3	Urban	5	
4	Rural	2	
…..			
1,000	Rural	1	

Samples to assess the level of EFC and calibrate the risk-based tools do not need to be substantial to be use

3. Do inspections on the random sample of beneficiaries

Sample of beneficiaries

No	Area	No of members	etc.	Result of the investigation
1	Rural	1		Non-fraud
2	Rural	4		Non-fraud
3	Urban	5		Fraud
4	Rural	2		Non-Fraud
…..	…	…		…
1,000	Rural	1		Fraud

+ Input the results of the social investigation in the database of sampled beneficiaries.

4. Identify the cases with most probability of EFC based on their characteristics

The simplest method: Tables

	% in total population	% of fraudsters
Urban	70	6
Rural	30	20
Total	100	10

No of members	% in total population	% of fraudsters
1	10	2
2	40	5
3	30	10
4+	20	25
Total	100	10

	No of members	% in total population	% of fraudsters
Urban	1	7	1
	2	30	3
	3	25	8
	4+	8	**14**
Rural	1	3	5
	2	10	10
	3	5	**20**
	4+	**12**	**32**
Total		100	10%

5. Identify in the total population the cases having the characteristics that flag a higher risk of fraud

Population of beneficiaries

No	Area	No of members	etc.
1	Urban	2	
2	**Rural**	**3**	
3	Rural	1	
4	**Rural**	**4**	
5	Rural	7	
6	Urban	2	
7	**Urban**	**4**	
…..			
…..			
100,000	Rural	1	

6. Do inspections primarily on the cases with higher risk of EFCin the population.
7. Review the model based on new iterations.

Profiles are not always so clear-cut…

Most of the times, the available variables do not predict clearly the beneficiaries committing fraud

We need to use more complex statistical techniques and more characteristics of the beneficiaries to predict the probability of EFC

Examples of techniques:

- Classification trees
- Logistic regressions
- Linear regressions
- Cluster analysis

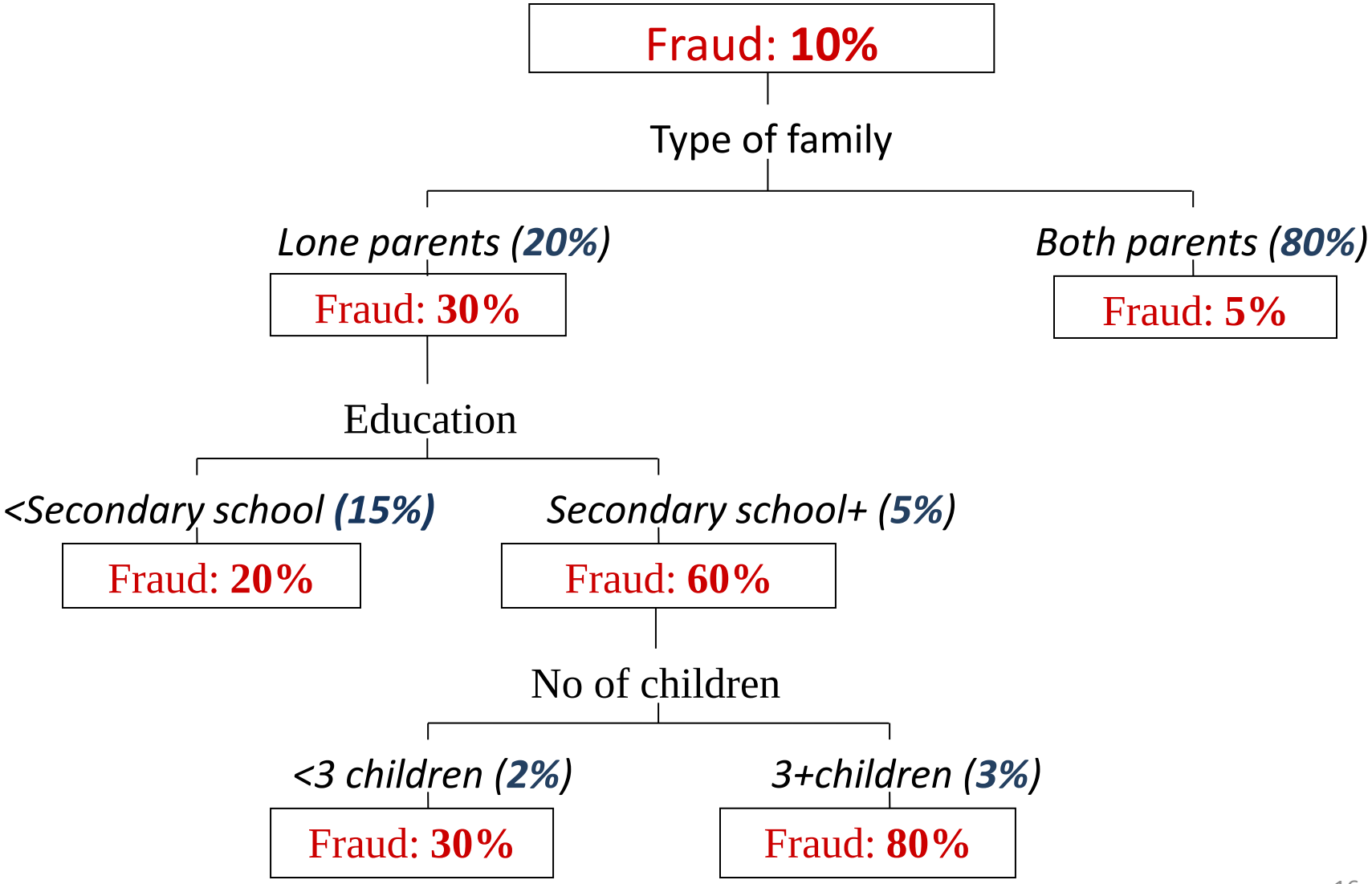
Classification trees

Automatic procedures exploring and analyzing data to identify the characteristics that are more likely to predict the beneficiaries committing fraud

No prior model specification of explanatory variables is needed

The technique identifies groups of individuals that are as homogeneous as possible based on a set of predefined variables.

Example of a Classification tree



Logistic regression

It identifies the key beneficiary characteristics that contribute to whether or not a case is fraudulent or in error.

Weights each characteristics according to its importance in identifying irregularity to provide an overall risk score for each case.

The risk score can take a value of between 0 and 1 (with 0 not at all likely to be in error and 1 more likely to be in error)

Example of a logistic regression model

	Odds Ratio (Fraud *versus* Non-Fraud)
"Lone parent" *versus* "Both parents"	2.5
"Rural" *versus* "Urban"	1.5
No of members in the household	1.25
"Secondary school" *versus* "No school/primary)"	3
"Tertiary school" *versus* "No school/primary)"	5

A lone parent beneficiary has 2.5 times higher odds to fraud than beneficiary family with both parents, holding all other factors constant.

Example of a logistic regression model

Population of beneficiaries

No	Family	Area	No of members	Education	Probability of fraud	Risk of fraud
1	Lone parent	Rural	5	Tertiary	0.9	**High**
2	Both parents	Urban	3	No education	0.1	Low
3	Lone parent	Rural	2	No education	0.3	Low
4	Both parents	Rural	5	Tertiary	0.7	**High**
5	Both parents	Rural	2	Secondary	0.5	Medium
…..						
100,000		Rural	1			

Outline

- **Why** inspections based on risk-profiling
- **How** to do inspections based on risk-profiling
- **Practical aspects** of doing inspections on risk-profiling
- **Plans** for risk-profiling in Romania

Needed resources

- Risk-based inspections can be implemented successfully even in environments with limited technical resources.
- Database with beneficiaries and their characteristics
- Social inspectors using the results of the risk-profiles and provide input in the feedback loop
- A system to input the results of the social inspections in the database of the beneficiaries' sample
- Statistical team of 3-5 persons experienced with data management, sampling techniques, and inferential statistics

Outline

- **Why** to do inspections based on risk-profiling
- **How** to do inspections based on risk-profiling
- **Practical aspects** of doing inspections on risk-profiling
- **Plans** for risk-profiling in Romania

Scheduled activities

- Individual consultancy to improve the technical skills of the statistical team
- Consultancy to undertake and improve the first rounds of social inspections based on risk profiles
- Increase the quality of available data in SAFIR (data auditing)
- Input/clean the data from the previous campaigns on random samples for Family Allowances and Child Raising Benefits
- Inspections based on risk-profiling for five benefits:
  - Guaranteed Minimum Income
  - Family Allowances
  - Heating Benefits
  - Child Raising Benefits
  - Disability allowances

Individual consultancy to improve the technical skills of the statistical team

"The objective of the consultancy is to prepare the participants to use STATA for the analyses required by their jobs. Specifically, the consultant has to:

a. *Train the Ministry's team to be able to apply the statistical methods that are generally used for predicting error and fraud*
b. *Help the Ministry's team understand the logic of selecting the cases for the next campaigns based on risk profiles*
c. *Train the Ministry's team to select samples for social inspection campaigns for both the identification of the groups with highest risks of fraud and error and for the fine-tuning of the next risk profiles"*

Consultancy to undertake and improve the first rounds of social inspections based on risk profiles

Specifically, the consultant is expected to undertake the following activities:

- To match the data available from previous inspection campaigns with the variables from other datasets that may be relevant for building more precise risk profiles
- To identify the variables that might be useful for building risk profiles that are not collected properly at the moment but could be adequately collected in the future;
- To build initial risk profiles for the before mentioned benefits using alternative statistical methods (i.e. logistic/linear regression, discriminant analysis, decision trees, neural networks, cluster analysis) in order to identify the best ways of predicting fraud and error

Consultancy to undertake and improve the first rounds of social inspections based on risk profiles

"…

- Do qualitative interviews with social inspectors to identify new possible risk factors that were not taken into account in the first quantitative analysis
- Select the cases for the next campaigns based on risk profiles (for each of the five benefits)
- Assist the team to build /refine the templates in which to input the data from social inspections for the five benefits in order to collect more reliable data
- Fine-tune the risk profiles based on the newly collected data and train the team on how to use the new risk profiles and do fine-tuning in the future"

Plan for social inspections based on risk-profiles

Putting in place the statistical team
STATA acquisition
Training for Statistical team
TORs finalized
Training package finalized
List of trainees confirmed
Training delivered
Assessment of the human resources for the inspection campaigns in 2014/2015
Estimation of the number of inspectors available for campaigns per semester
Estimation of the duration of each of the campaigns (in months)
Estimation of the number of inspections on a. disability benefits, b. family allowances, c. guaranteed minimum income, d. heating benefit, e. child raising allowances that can be conducted by an inspector on average per day

Plan for Family Allowances' Inspections

Risk profile analysis for FA

- Data cleaning of the information from the campaign from 2011
- First risk profile analysis

Campaign based on the first risk profile

- Write-up of the methodology for future campaigns (steps to be undertaken, refinement of protocols for social inspectors, design of the methodology for the collection of the results of social inspections etc.)
- Sample selection for the campaign based on the first risk profile
- Inspection Campaign
- Feedback from the field on the methodology of the campaign (to improve the procedures for the future campaigns)
- Data input (of the campaign's results)

Second Risk profile analysis for FA (improvement of the first Risk profile)

Campaign based on the second risk profile

- Sample selection for the campaign based on the second risk profile
- Inspection Campaign
- Data Input(of the campaign's results)

Plan for Heating Benefits' Inspections

Campaign based on a random sample

- Write-up of the methodology for future campaigns (steps to be undertaken, refinement of protocols for social inspectors, design of the methodology for the collection of the results of social inspections etc.)
- Random sample selection for the campaign
- Inspection Campaign
- Feedback from the field on the methodology of the campaign (to improve the procedures for the future campaigns)
- Input of data (of the results of campaigns)

First Risk profile analysis for HB

Campaign based on the first risk profile

- Sample selection for the next campaign based on the first risk profile
- Inspection Campaign
- Data Input (of the campaign's results)

Second Risk profile analysis for HB (improvement of the first Risk profile)