SWEDISH SOCIAL INSURANCE INSPECTORATE

Audit strategy for temporary parental benefit

Leif Appelgren

Working Paper 2012:2

Audit strategy for temporary parental benefit

Leif Appelgren,
Economic Information Systems,
Department of Management and Engineering
Linköping University

Abstract

The aim of this project has been to study the possibility to apply audit strategies developed for taxation on fraud and involuntary errors in the social benefit sector. The efficiency of different audit strategies is compared using a computer-based optimization algorithm.

Two types of audit strategies are used in this study. One is to adapt the audit intensity to the propensity for errors and fraud in different segments of the groups studied. The other type of audit strategy is based on adaptation of behaviour through information concerning the audit intensity. A model for determination of optimal tax audit strategies of the latter type was developed by Erard & Feinstein in 1994.

This study is based on data from a large study of temporary parental benefit performed by the Institute for Evaluation of Labour Market and Education Policy (*Institutet för arbetsmarknadspolitisk utvärdering, IFAU*) in 2006.

The study has shown that it is possible to apply the Erard & Feinstein model on benefit fraud. However, the solution method developed by Erard & Feinstein has proven to be non-optimal. A new solution method based on simulation has been developed and used in the study.

Content

Summary

The aim of this project is to study the possibility to apply audit strategies developed for taxation on fraud and involuntary errors in the social benefit sector. The term fraud refers to voluntary errors in the submission of information from the recipient of benefits. The efficiency of different audit strategies is compared using a computer-based optimization algorithm.

The term audit strategies refers to methods to audit and verify the data which individuals and corporations supply as basis for decisions regarding taxes or benefits. The aim is to find methods which are more efficient than random or total audits of the group studied. A more efficient method has less remaining errors and fraud with a given audit cost, alternatively is less costly with a given level of audit intensity.

Two types of audit strategies are used in this study. One is to adapt the audit intensity to the propensity for errors and fraud in different segments of the group studied. The propensity for errors and fraud is estimated from earlier audits.

The other type of audit strategy is based on adaptation of behaviour through information. Auditing has two effects, the direct effect of discovery and sanctions, and the indirect, preventive effect that the auditees adapt their fraudulent behaviour to the risk of being audited. The indirect effect is difficult to model, but analytical models can be developed under the assumption that the auditees act rationally.

The type of model used here is based on that the audit intensity is controlled by a variable and that the auditees are informed of this relation between control variable and audit intensity. For income tax purposes, the declared income normally is the control variable, such that those who declare the lowest income are subject to the highest audit intensity. The rational taxpayer understands that he/she by reducing the amount of fraud and thus increasing the declared income reduces the risk of being audited. A model for determination of optimal audit strategies was developed by Erard & Feinstein in 1994 (the E&F model). This model allows a separation of the studied group in inherently ”honest” auditees always declaring their true income, and ”rational fraudsters”. In order to apply this model, knowledge is required regarding the distribution of true income and the portion of inherently honest auditees.

In an application on fraud in the social benefit sector, the amount of benefits claimed can be used as the control variable. In the case studied on temporary parental benefits (Tillfällig föräldrapenning, TFP), it is practical to use the amount of benefits claimed during a certain period (the control period) as the control variable. Those who make large and/or frequent benefit claims will then risk a higher audit intensity.

This study is based on data from a large study of TFP performed by the Institute for Evaluation of Labour Market and Education Policy (*Institutet för arbetsmarknadspolitisk utvärdering, IFAU)* in 2006. TFP is administered by the Swedish Social Insurance Agency *(Försäkringskassan, FK)*, which supplied the data to IFAU. In the IFAU study, more than 2000 persons were audited, resulting in that errors were found in 16% of the audits. All TFP benefit payments to the audited persons during a two month period (*the audit period*) were audited. The database also included all payments made to the audited persons during an earlier five month period (*the reference period*) and the time between the reference period and the audit period.

The analysis reported in this paper refers to a period of about 8.5 months, including the audit period, the reference period and the one month period in between. As no audits were made outside the audit period, it has been assumed that each individual has the same relative error amount as during the two month audit period.

Those who were audited in the IFAU study were divided into three groups A, B and C, where group A received a warning letter (”You have been selected for special scrutiny…”) plus an information letter concerning the regulations regarding right to TFP. Group B received the information letter only whereas group C received the warning letter only. No audits were made among those who did not receive any letter (group D).

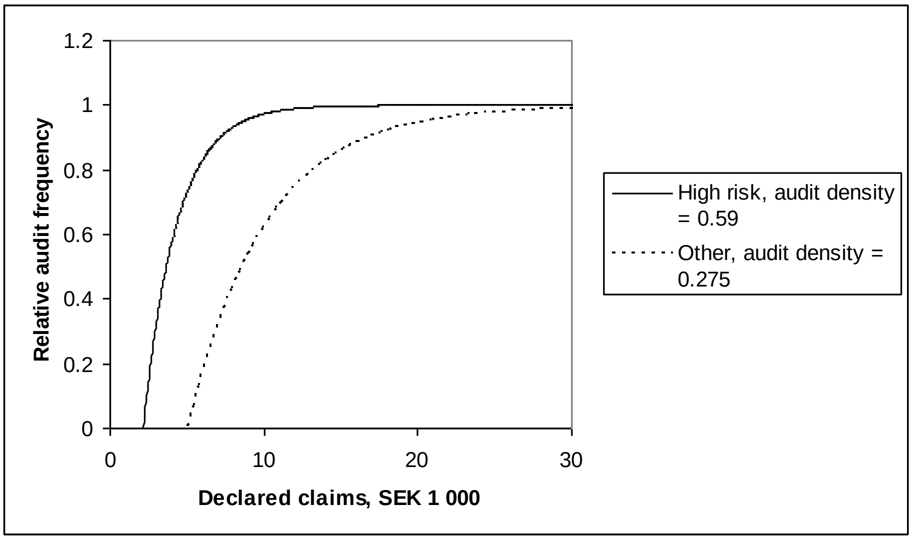
An important parameter in the model used is the fraction of inherently honest persons. This is best estimated from data for those who were least influenced by the two letters. The study is therefore mainly based on the audits in group B. However, all three groups have been used for segmentation into risk segments as the data material otherwise would have been too small for a statistical analysis.

Information is recorded in the IFAU database regarding personal data such as sex, place of residence, employment sector, age, education and income. As the database is not perfectly consistent, the population studied consists of audited persons for whom personal data and claims data for the 8.5 month period are available.

The segmentation is based only on the portion of persons with errors and not on the size of those errors. The segments *Stockholm region, Income below SEK 100 000, Education not reported*, *Sector not reported* and *3-4 children* are referred to a *High risk segment.* The remaining population is denoted *Other*. *High risk* means that the probability of errors is high.

In this paper, *audit frequency* denotes the audit intensity as a function of the amount of benefits claimed, whereas *audit density* is the portion of persons audited in a group or segment.

The figure below shows the typical form of optimal audit frequency functions, where no audits are made on persons with low benefit claims whereas all persons with very high claims are audited.



Efficient auditing implies that the audit intensity is higher in segments where the error probability is high and vice versa. Persons in the high risk segment are audited much more frequently than persons in the low risk segment.

The behaviour of the auditees who act rationally is assumed to be affected by the size of sanctions in case they are discovered. In the Swedish tax system, there is a tax surcharge which is proportional to the tax evaded, but in the social benefit sector, a corresponding sanction has not yet been introduced (benefit surcharge).

This study mainly treats a Base case without sanctions, but in addition a case with a surcharge amounting to 25 % of the fraudulent claim is studied. Such a sanction has a substantial positive effect on the audit efficiency.

In the E&F model used here, it is assumed that all errors are voluntary, i.e. fraudulent. The commission from ISF included an attempt to extend the model to handle also involuntary errors, since FK estimates that approximately 50 % of the errors are involuntary. The effect of involuntary errors is that the audits become less efficient since there will be no adaptation of behaviour from those who make involuntary errors. As mentioned above, it is a prerequisite for the E&F model that the audit strategy is announced in advance such that the rational fraudster can adapt his/her behaviour.

The E&F model has been developed further in this study in order to include a mix of fraud and involuntary errors. It has also been extended to handle probability distributions of true claims which are partly continuous and partly discrete.

The optimization of audit strategy can be made either with a fixed amount of audit resources or with a fixed cost per audit. Results are reported for both these cases with a emphasis on the latter case. FK has reported its internal time and cost estimates for audits made over telephone. The unit audit cost is estimated to SEK 150-188.

A few interviews have been made with parents, employers, schools and pre-schools regarding the resources required for telephone audits. The total audit cost for FK, school/pre-school and employer is estimated at SEK 280-580. The parents are not involved in a system with telephone audits, instead they carry a burden in the present system with the administration of a certificate regarding absence from school/pre-school.

The main results are condensed in the table below. *Total cost* is the sum of audit cost (taking into account the audit density) and the average cost for remaining fraud and errors. The table illustrates what has been stated above, i.e. that the total cost is reduced when a sanction cost is introduced and that the total cost is increased when the errors are partly involuntary.

Total cost per person in the base case is SEK 130-505, dependent on whether external costs are included or nor. The annual cost per person is 12/8.5 higher, i.e. SEK 184-713. As 683 000 persons made TFP claims in 2006, the annual cost would be SEK 125-485 million.

	Internal costs only *SEK 150-188*		*Internal plus external costs* *SEK 280-580*	
	Optimal audit density	Total cost per person	Optimal audit density	Total cost per person
Base case	87 %	SEK 130-165	87 %	SEK 245-505
Segmentation	87 %	SEK 130-165	87 %	SEK 245-505
25 % sanction factor	70 %	SEK 105-130	70 %	SEK 195-405
50 % involuntary errors	93 %	SEK 140-175	93 %	SEK 260-540
Combination case	75 %	SEK 115-140	75 %	SEK 210-435

The TFP payments amounted to approximately SEK 3 900 million in 2006. The total cost for audits and remaining errors should in the base case vary between 3 and 12 % of the payments, which seems to be a realistic share.

The main conclusions of the study are

- It is possible to apply the Erard & Feinstein model on benefit fraud.
- The solution method developed by Erard & Feinstein has proven to be non-optimal. A new solution method based on simulation has been developed.
- A model for handling involuntary errors has been developed.
- It is possible to compare audit strategies according to this study with other audit systems, for instance the certificate system presently used in schools and pre-schools as a prerequisite for TFP payments.

1 Introduction

The Swedish Social Insurance Inspectorate (ISF) commissioned in January 2010 Linköping University to study the possibility to apply existing models for tax auditing on benefit fraud.

The tax audit models in question are based on that the fraudulent taxpayer males a rational choice of the amount of fraud in order to maximize his/her expected gain. The auditor optimizes its audit strategy in order to minimize the expected tax loss.

Gary Becker published in 1968 a criminological model which assumes that criminals make rational decisions regarding crime based on utility, risk for discovery and sanction and the cost of the sanction. This model was applied to tax fraud by Allingham & Sandmo (1972), Reinganum & Wilde (1986) and Erard & Feinstein (1994). Allingham & Sandmo (A&S) studied the optimal behaviour of taxpayers with a known audit cost and audit frequency and a quadratic utility function. Reinganum & Wilde (R&W) extended the model to study the optimal behaviour of the auditor, using an audit frequency varying with declared income. The R&W model uses a known audit unit cost and a linear utility function.

Erard & Feinstein (E&F) adapted the R&W model to optimization of audit strategy under given audit resources instead of a given audit unit cost. They also extended the model to comprise two categories of taxpayers, one with members which behave as rational criminals according to the Becker model and one with members which always declare their true income. This extension complicated the mathematics considerably.

The models above are presented in a survey paper by Andreoni et al (1998).

It should be remarked that none of the models above studied the effect of involuntary errors.

It is assumed in all models above that the fraudster makes a rational analysis and optimization of the fraud amount, taking into account the audit frequency in the A&S model and the decreasing audit frequency function in the R&W and E&F models. It is of course unrealistic to believe that the majority of fraudsters behave so rationally, but it is quite realistic to assume that fraudsters are affected by the knowledge that the amount of fraud has an effect on the risk of discovery. Studies made by Blumenthal et al (2001), Hasseldine et al (2007) and Appelgren (2008) show that information regarding audit strategy has an effect on taxpayer behaviour.

The aim of the project is to study the possibility to apply the E&F model on fraud and involuntary errors in the social benefit area.

ISF elected to carry out the project on Temporary Parental Benefit for Child Care (TFP) because a comprehensive data base was available from a study made in 2005-2006 by the Institute for Evaluation of Labour Market and Education Policy (IFAU) (Engström et al (2006)). TFP was paid out in 2006 to about 683 000 persons with about SEK 3 900 million.

It was apparent in the IFAU study that information regarding increased auditing had a marked effect on the errors, which makes it probable that the E&F model may be useful for design of audits of TFP and possibly also other social benefits.

When the E&F model is applied to tax fraud, the audit intensity is varied as a function of declared income. The optimal audit frequency declines with increasing income which gives the fraudster an incentive to cheat less in order to reduce the risk of discovery. It has proved suitable to let the claims of TFP during a certain period control the audit frequency, such that the fraudsters get an incentive to reduce the benefit claims, i.e. to cheat less in order to reduce the risk of discovery.

According to the Swedish Social Insurance Agency (FK) which administers TFP, voluntary errors represent only about 50 % of the errors which are discovered in the audits. An important part of the project has therefore been to attempt to extend the model to comprise both fraud and involuntary errors.

This working paper is an updated version of a previous working paper written in Swedish (Appelgren (2011)), also summarized in Molander (2011). The substantial difference is that a new and improved solution method is used in the present paper. The main results are quite similar.

In the paper, consistency in the terminology is sought in the following aspects:

- *Audit period* refers to the two month period in which claims were audited in the IFAU study
- *Claims* or *declared claims* refer to payments made during the total period, measured in SEK
- *Control period* refers to the period in which the total claims is used as the control variable for determination of audit frequency
- *Errors* refers to the errors discovered in the IFAU audits. *Error* also refers to the sum of fraud and involuntary errors in the extended model

- *FK* is an abbreviation of *Försäkringskassan (The Swedish Social Insurance Agency)*
- *Fraud* refers to voluntary errors in the supply of information from the claimant of benefits. It is assumed in the Erard & Feinstein model that all errors are fraudulent
- *Group* are the groups in the IFAU study, i.e. A, B, C and D
- *IFAU* is an abbreviation of *Institutet för arbetsmarknadspolitisk utvärdering (The Institute for Evaluation of Labour Market and Education Policy)*
- *ISF* is an abbreviation of *Inspektionen för socialförsäkringen (The Swedish Social Insurance Inspectorate)*
- *Reference period* refers to the five month period used in the IFAU study for determination of claim amounts unaffected by warning or information letters
- *Segment* refer to the subdivision of the groups according to the variables available in the data base
- *TFP* is an abbreviation of *Tillfällig föräldrapenning (Temporary Parental Benefit)*
- *Total period* refers to the period used as the control period in this study, approximately 8.5 months, consisting of the reference period (five months) and the audit period (two months) used in the IFAU study plus the time interval between the two periods
- *True claims* refer to declared claims less measured/estimated errors, i.e. correct claims according to regulations, measured in SEK

2 Data

Temporary Parental Benefit (TFP) is available to carers (parents or other persons) for care of sick children. It is also available for some other purposes of less importance, not included in the IFAU study. TFP can be paid out for entire days or part of a day. Overutilization may depend on fraud, involuntary errors of the carer and errors committed by the FK staff. The aim of the IFAU study was to measure the size of overutilization, measured in monetary terms or net days. Net days means that claims of parts of days are converted to claims of entire days.

The IFAU measurement was largely based on the effect of information to the carers. After a *reference period* when the claims of TFP were measured, a warning letter was distributed, after which a new measurement of claims was made during the *audit period*. The message of the warning letter was: ”You have been selected for special scrutiny…”. An information letter was also used in the study, describing the regulations concerning TFP.

The reduction of claims between the reference period and the audit period was used as a measure of fraud. In addition, audits were made during the audit period in order to measure remaining errors. The audits were made by telephone calls to schools/pre-schools and employers. Frequent errors were that the carer had been working or that the child had been in school/pre-school during the claims period.

The population was divided into four groups:

A: Both warning letter and information letter

B: Information letter only

C: Warning letter only

D: No letter

Audits were carried out in groups A-C, but not in the unaffected group D.

The database used in the IFAU study as well as the present study consists of four data files received from FK. Additional details are available in the Appendix.

The Audit file contains data for about 2 400 audits of TFP payments carried out during the *audit period* March 29 – May 31, 2006. The most important data are id-number, amount paid, Right/Wrong and error amount.

The Payment file contains all TFP payments nationwide during the *reference period* October 1, 2005 - February 28, 2006, the *audit period* plus the time between the two periods, together called the *total period*. The file contains also some payments outside the total period, such that the effective length of the period covered is approximately 8.5 months with about 2.1 million payments. Most important data are id number for adult, id number for child and amount claimed.

The Population file contains data concerning the 1.3 million persons who were eligible for TFP claims during the period studied. The file contains data on sex, time of birth, place of residence, employment sector, education and income.

The Relation file contains data regarding the 1.1 million children who were the basis for TFP claims during the period studied. The file contains id numbers for maximally six related persons.

The data required for application of the E&F model are the true probability distribution of the control variable and the portion of non-fraudulent persons.

The total claim amount of TFP during the total period was selected as the control variable. Fraudulent persons increase their claims, thus it is logical to impose a higher audit intensity for persons with large claims. If such a strategy is announced, the fraudulent persons have an incentive to reduce their amount of fraud in order to reduce the risk of discovery.

It would have been desirable to carry out the analysis in this study on a group not affected by the two letters, i.e, group D, but no audits were made in this group in the IFAU study. The analysis below is therefore mainly based on the audits in group B, which is affected by the information letter only. Regrettably, the number of audited individuals in group B was only 339 compared to 1 271 in group A and 356 in group C. Data from all three groups have therefore been used for the segmentation in order to reduce the statistical uncertainty. The use of data from groups A and C is based on the hypothesis that all segments are influenced in the same manner by the two letters.

One motive for the warning letter in the IFAU study was to measure errors and fraud which cannot be detected in an ordinary audit. One example is "care of healthy child", where the carer is absent from work and the child is absent from school/pre-school. Motives for this kind of fraud may be that the carer is sick or wishes to spend time together with the child.

In this study, the effect of the warning letter is not taken into account. The analysis is based entirely on the audits carried out during the audit period.

All payments during the audit period for the audited persons are registered in the Audit file. All payments which an audited person has received during the audit period are audited, thus the number of audits is larger than the number of audited persons. The analysis in this paper is based on total payment and total error amount for the audited persons.

The sample used consists of the audited persons, except a) those who have a zero total payment during the audit period (17 persons) and b) those persons who are missing in the Payment file (64 persons). For four persons, the total error amount exceeds total payments. In these cases, the error amount has been reduced to the payment amount.

Table 1. Characteristics for groups A, B and C

	Group A Both letters	*Group B Information letter*	*Group C Warning letter*	*Groups ABC*
Number of audits	1 574	427	447	2 448
Number of persons	1 271	339	356	1 966
Number of persons with zero payments	12	3	2	17
Number of persons missing in the Payment file	31	15	18	64
Net number of persons	1 228	321	336	1 885
Number of persons with errors	208	68	36	312
Fraction of persons with errors	16.9 %	21.2 %	10.7 %	16.6 %
Number of persons with true claims > 0	1 108	285	316	1 709
Number of erring persons with true claims > 0	88	32	16	136
Fraction of erring persons with true claims > 0	7.9 %	11.23 %	5.1 %	7.96 %
Data for audit period				
Total claims, SEK 1 000	1 832	492	515	2 839
Error amount, SEK 1 000	212	74	33	319
Average error per erring person, SEK	1 019	1 084	927	1 022
True claims, SEK 1 000	1 621	418	482	2 520
Error share of total claims	11.6 %	15.0 %	6.5 %	11.2 %
Average total claims per person, SEK	1 492	1 531	1 532	1 506
Average true claims per person, SEK	1 318	1 302	1 433	1 337
Maximum true claims, SEK	11 400	7 722	10 388	11 400

Data for total period				
Total claims, SEK 1 000	8 680	2 311	2 478	13 469
Estimated error amount, SEK 1 000	1 107	392	214	1 713
Average error amount per erring person, SEK	5 324	5 759	5 951	5 491
True claims, SEK 1 000	7 572	1 920	2 264	11 756
Error share of total claims	12.8 %	16.9 %	8.64 %	12.7 %
Average total claims per person, SEK	7 068	7 200	7 375	7 145
Average true claims per person, SEK	6 166	5 980	6 738	6 237
Maximum true claims, SEK	67 314	40 117	39 516	67 314
Number of claims	7 221	1 894	2 052	11 167
Average number of claims per person	5.88	5.90	6.11	5.92

Out of the 68 erring persons in group B, 36 have true claims = 0. Thus, 285 persons have true claims > 0, of which 32 have made errors, i.e. 11.2 %. This is a more correct measure of the share of erring persons compared to the erring share of all persons in group B, where the erring share is 21.2 %. The reason is that the persons who have true claims = 0 and no errors are not included in the audited group. In the analysis below, the sample is extended with a number of non-erring persons with true claims =0.

Characteristic data for the three groups are shown in Table 1. As errors have been measured during the audit period only, it is assumed that each individual make the same relative error during the total period as during the audit period. This assumption is most plausible for group B, but not entirely logical for groups A and C since they have been affected by the warning letter sent out between the reference period and the audit period. This is further discussed in Section 9.2.

The error share of total claims in group B increased to 16.9 % for the total period compared to 15 % for the audit period. This is explained by the fact that the erring persons make relatively larger claims during the total period than the non-erring persons.

683 000 persons claimed TFP during 2006, with a total amount of about SEK 3 900 million. With 21.2 % erring persons, 145 000 persons should have made an error during the year, at a cost of 16.9 % of the total claims, i.e. approximately SEK 660 million.

A histogram for the distribution of true claims >0 is shown in Figure 1, which shows the similarity between group B and the compounded data (group ABC). Please note that the class intervals on the x axis have a varying width**.** See also Figure 2 in Section 3.

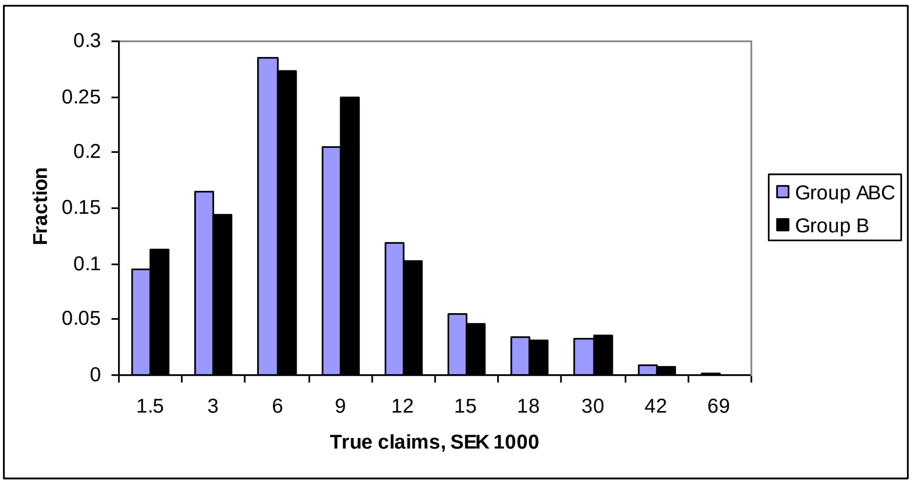


Figure 1. The distribution of positive true claims in group B and groups ABC for the total period. The numbers on the x axis refer to the upper class limit

3 Segmentation

General

The segmentation is made on the compounded data for the three groups, denoted group ABC, and only for persons with true claims >0, a total of 1 709 persons according to Table 1. In Table 2, results of the segmentation are shown for the variables sex, education, employment sector, place of residence, age, income and number of children. It should be remarked that all variables except sex and number of children have a finer classification than what is shown here. The choice of segments is therefore arbitrary to some extent.

The selection of segments into risk segments is made by a simple statistical test.

For the share of erring persons, the standard deviation is $\sigma = \sqrt{(p(1-p)/n)}$, where p is the true share of erring persons and n the number of persons in a specific segment. The true share of erring persons is approximated with the average for all persons with true claims > 0, i.e. $p = 0.0796$. Segments with significant ratio $\Delta m/\sigma$ are marked with asterisks in Table 2.

Table 2. Segmentation of groups ABC

	Number with true claims > 0	*Number of erring persons with true claims > 0*	*Share of persons with errors*	*Δm*	*σ*	*Δm/σ*
Sex						
Men	642	43	0.0670	-0.0126	0.0107	-1.18
Women	1067	93	0.0872	0.0076	0.0083	0.92
Total	1709	136	0.0796			
Education						
Comprehensive school	124	14	0.1129	0.0333	0.0243	0.73
Secondary school	936	79	0.0844	0.0088	0.0048	0.55
University	644	41	0.0637	-0.0159	0.0107	-1.49
Not reported	5	2	0.4000	0.3204	0.1210	2.65**
Total	1709	136	0.0796			

Employment sector						
Private	1042	88	0.0845	0.0049	0.0084	0.58
City	410	30	0.0732	-0.0064	0.0134	-0.48
County	122	5	0.0410	-0.0386	0.0245	-1.58
State	79	4	0.0506	-0.0289	0.0304	-0.95
Total public sector	1653	127	0.0845	0.0049	0.0084	0.58
Not reported	56	9	0.1607	0.0811	0.0362	2.24*
Total	1709	136	0.0796			
Place of residence						
Stockholm county	402	46	0.1144	0.0348	0.0135	2.58**
Greater Gothenburg region	193	19	0.0984	0.0189	0.0195	0.97
Greater Malmö region	125	7	0.0560	-0.0236	0.0242	-0.97
Other	988	64	0.0648	-0.0148	0.0086	-1.72
Not reported	1	0	0.0000	-0.0796	0.2706	-0.29
Total	1709	136	0.0796			
Age						
Born 1943-61	148	11	0.0743	-0.0053	0.0222	-0.24
Born 1962-76	1434	109	0.0760	-0.0036	0.0071	-0.50
Born 1977-85	127	16	0.1260	0.0464	0.0240	1.93*
Total	1709	136	0.0796			
Income, SEK 1000						
0-100	77	13	0.1688	0.0893	0.0308	2.89**
100-200	578	49	0.0848	0.0052	0.0113	0.46
200-300	734	51	0.0695	-0.0101	0.0100	-1.01
300-400	237	18	0.0759	-0.0036	0.0176	-0.21
>400	83	5	0.0602	-0.0193	0.0297	-0.65
Total	1709	136	0.0796			
Number of children						
One child	907	62	0.0684	-0.0112	0.0090	-1.25
Two children	725	62	0.0855	0.0059	0.0101	0.59
3-4 children	77	12	0.1558	0.0763	0.0308	2.47*
Total	1709	136	0.0796			

* <.05, **<.01

Segmentation variables

Sex

Women have a 2 % higher erring share than *Men*, but the difference is not significant.

Education

The segment comprehensive school (grundskola) has a larger erring share than university and secondary school (gymnasium), but not significantly. Five persons lack education data. This small segment has a significantly high erring share.

Employment sector

For 56 persons, *no employment sector* is reported. This segment has a significantly high erring share.

The segments *Employment in state and county administration* have low erring shares, but not significantly so.

Place of residence

Stockholm county has a significantly high erring share. The *Gothenburg region* is slightly above average whereas the *Malmoe region* and the rest of Sweden is slightly below.

Age

The selection of age classes is based on earlier studies of the data, where an increased share of erring persons was observed for the youngest persons (below 29 years) and the oldest persons (above 45 years). A high share of erring persons is noted for the youngest segment *Born 1977-85* close to the significant level. This segment has been included in the *High risk* segment.

Income

In the data base, four different income data are included in the data base, but neither FK nor IFAU has been able to supply a definition. Therefore, the sum of the four income items has been used as the income variable. The segment SEK 0-100 000 shows a significantly high share of erring persons.

Family situation, number of children

It is not possible to ascertain from the data whether a child lives with two carers at the same address. Instead, it is possible to state if the child has one carer only or if it has multiple carers living in different parishes. As this analysis is technically complicated and still cannot give a certain answer whether the child lives in a nuclear family, the nuclear family variable is not included in the study.

The number of children has been approximated with the number of children with the same carer for which claims have been made during the total period. This is of course an underestimate of the true number, but this is considered to be not essential for the analysis. The segment *Four children* consists of two persons only, it has therefore been merged with the segment *Three children*.

Table 2 shows that the share of erring persons increases with the number of children and that the difference is significant for the segment *3-4 children*.

Risk segmentation

According to the test in Table 2, *Education not reported, Employment sector not reported*, *Stockholm county, Born after 1976, Income below SEK 100 000* and *3-4 children* should be referred to the High risk segment. No segment has a significantly lower share of erring persons than the average, thus the segmentation results in two risk segments only, *High risk* and *Other*, see Table 3. The share of persons with errors is 12.1 and 5.5 %, respectively.

Table 3. The two risk segments for group ABC

	Number with true claims > 0	*Number of erring persons with claims > 0*	*Share of erring persons*	*Δm*	*σ*	*Δm/σ*
High risk	626	76	0.1214	0.0418	0.0108	3.87
Other	1083	60	0.0554	-0.0242	0.0082	-2.94
Total	1709	136	0.0796			

Adjustment of data for group B

The analysis is based on the assumption that group B is only slightly influenced by the information letter and that data for group B therefore can be considered representative for an unaffected population (group D).

The share of persons with errors is an important parameter in the analysis together with the distribution of true claims. In the analysis without segmentation, the measured share of erring persons in group B is used. The erring share in the two risk segments are calculated from the total population, while it is apparent from Table 1 that the erring share is considerably higher for group B. The erring share in the two risk segment is therefore adjusted upwards for the analysis of group B.

The share of erring persons in group B amounts to 11.23 % whereas for the total material it is 7.96 %. The adjustment is made with the factor 0.1123/0.07965 = 1.411.

Table 4. Adjustment of erring share for group B

Erring share, %	*Group ABC*	*Factor*	*Group B*
All	7.96	1.411	11.23
High risk	12.14	1.411	17.13
Other	5.54	1.411	7.82

It can be discussed whether the distribution of true claims should be calculated from group B only or from all three groups. In an analysis of group B unsegmented, it seems natural to use the claims distribution for group B. Given that the segmentation is based on data from all three groups, it would be natural to use the distribution for all three groups in a segmented analysis. It is however important that it shall be possible to compare the segmented and the unsegmented analyses in a consistent manner; therefore the distribution for all three groups has been used in all analyses.

True claim distributions

The empirical distribution of true claims for group B and groups ABC is shown i Figure 2. The difference between the distributions is small.

The maximum true claim is SEK 67 300. The ABC distribution is subdivided in 10 classes according to Table 5.

Table 5: Classes in the true claim distribution

Interval, SEK 1 000	*Class width, SEK 1 000*	*Number of classes*
0-3	1.5	2
3-18	3	5
18-42	12	2
42-69	27	1

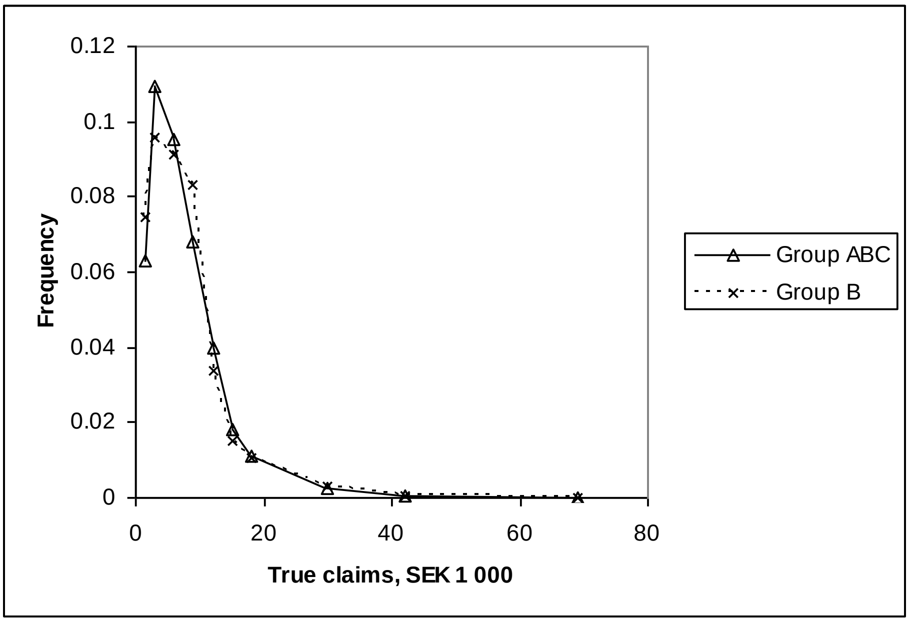


Figure 2. The true claims distribution for group B and groups ABC

The true claims distribution is extended with a discrete probability for true claims = 0, calculated below.

The number of persons with errors in groups ABC amounts to 312. Out of those, 176 have true claims = 0 and 136 have true claims > 0. Groups ABC contain 1 709 persons with true claims > 0. In order to obtain a correct distribution of erring persons between those with true claims = 0 and those with true claims > 0, one should add 1 709x176/136 = 2 212 persons with true claims = 0. Those persons make up 176/312 = 56.4 % of the claims distribution.

The reasoning above is based on the assumption required by the E&F model that all persons have the same propensity for errors.

Figure 3 shows the true claims distribution for the two segments. It is obvious that the high risk segment has less weight at the lower end and a higher weight for medium and high claims. The mean value of the High risk segment exceeds the mean of the entire sample with about SEK 1 000, which is a significant difference. A chi square test also reveals that the high risk distribution is significantly different from the ABC distribution.

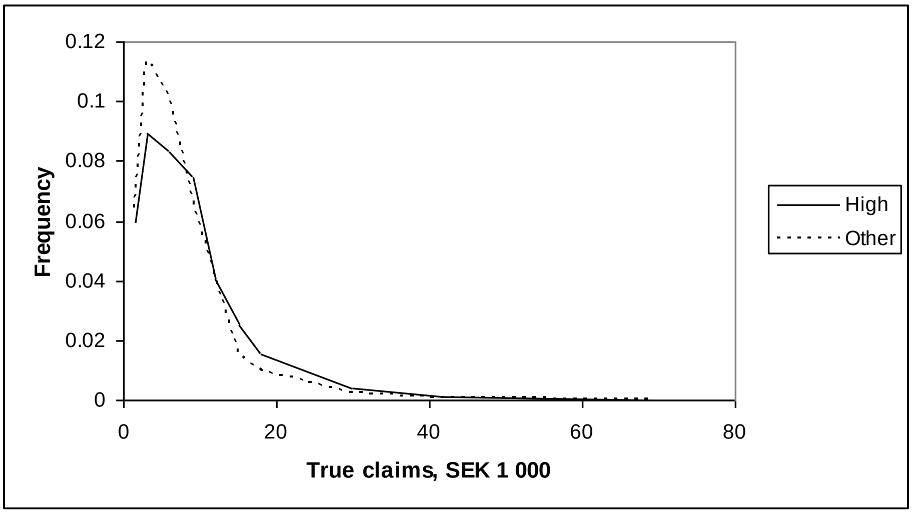


Figure 3. The true claims distribution for the segments *High risk* and *Other*

It should be noted that this is the distribution of true claims, corrected for errors and fraud. There is no simple explanation to the fact that the high risk group has a higher mean. A possible explanation is that the high risk segment consists of a higher share of persons with impaired health and thus a higher propensity for child sickness.

4 Optimal audit strategy

4.1 Audit method

For tax audits, it is natural to use annual income tax reports as the database, with the declared annual income as the control variable. For audits of TFP as well as other social benefits, it is natural to have a running audit program with a certain number of audits per week or month. What then shall be audited: Is it single large claims or all claims made during a certain period? In the latter case, how long should this period be?

The application of the E&F model on TFP means that the audit intensity should increase with the claim amount, with no audits at all for the smallest claims. We assume that the optimal strategy is to audit no claims of SEK 500 or less if audits are based on single claims. A person who makes one claim of SEK 2 000 during a certain period would then run a high risk of being audited, whereas a person making four claims of SEK 500 each would not be audited at all. The conclusion is that audits based on the size of individual claims should be discarded.

The length of the control period is an optimization problem in itself. It depends on the availability of claims data at employers and schools/pre-schools, the audit unit cost and the expected cost of errors and fraud. In this paper, results are reported for an 8.5 month control period. In an earlier paper (Appelgren (2011)), results for a two month control period are also reported. A conclusion in that paper was that the two month period was inferior to the 8.5 month period, it is therefore not treated here.

4.2 The Erard & Feinstein model

The analytical model for optimal tax audits stems from Reinganum & Wilde (1986a and b) and Erard & Feinstein (1994). It is based on the assumption that the fraudster behaves rationally and optimizes his/her economic gain by taking the risk for discovery and the size of sanctions into account. The optimal audit frequency declines with the declared income, giving the fraudster an incentive to increase the declared income in order to reduce the risk of discovery.

The model is modified for benefit fraud such that the audit frequency increases with the claims of the benefit in question. The fraudster, who is assumed to be aware if the audit strategy, gets an incentive to reduce the fraud amount.

The E&F model, adapted to benefit fraud, is based on the following assumptions:

- The sanction on discovery is proportional to the fraud amount. The sanction may be zero.
- The auditor uses no other information about the auditee than his/her claims of the benefit in question.
- The distribution of true claims (i.e. claims adjusted for errors and fraud) in the group is known, for example from earlier random audits.
- The group consists of two parts, where a known fraction is error free while the remaining fraction consists of ”rational fraudsters”.
- All fraud is discovered in an audit.

The model has been extended to handle a discrete component in the true claims distribution for true claims = 0. It has also been extended to handle a mix of fraud and involuntary errors.

Important terms

The audit frequency function p(x) refers to the audited share for a specific value of declared claims *x.*

Relative audit frequency is the ratio between audit frequency and the *Critical audit level 1/(1+a)*, where *a* is the ratio between sanction amount and fraud amount. For tax fraud, the standard Swedish sanction is a 40 % tax surcharge**.** Thus the Critical audit level is 1/1.4 = 0.71. Fraud is not worthwhile for a rational fraudster if the audit frequency exceeds the Critical audit level. The relative audit frequency will thus never exceed unity.

Audit density refers to the audited share of the group. The audit density is the average audit frequency in a group.

Relative audit density is the ratio between audit density and the *Critical audit level 1/(1+a).*

The model

The principles for the E&F model, adapted to benefit fraud, is that the fraudster maximizes his/her utility

$U = (x-y)(1-(1+a)p(x))$

where *y* is the amount of true claims, *x* the amount of declared claims and *p(x)* the audit frequency as a function of declared claims.

The auditor wishes to minimize the average net fraud amount, ”the error cost”, which is

$$C = (1-Q) \int (x-y)(1-(1+a)p(x(y)))f(y)dy$$

under the constraint

$$\int p(x)g(x)dx = B$$

where Q is the fraction of non-fraudsters, B the audit density, $f(y)$ the distribution of true claims and $g(x)$ the distribution of declared claims.

The constraint can be included in the objective function via a Lagrange multiplier λ. The objective function then becomes

$$C_{\lambda} = (1-Q) \int (x-y)(1-(1+a)p(x(y)))f(y)dy + \lambda \int p(x)g(x)dx$$

In case that the audit unit cost c is given externally instead of the audit density, the objective function becomes

$$C_{c} = (1-Q) \int (x-y)(1-(1+a)p(x(y)))f(y)dy + c \int p(x)g(x)dx$$

The two objective functions are identical if c is replaced with λ. The difference is that c is given externally whereas λ has to be adapted such that the constraint is satisfied. The formulation with a given audit unit cost c is natural in a flexible organization, but the audit constraint formulation may be required in cases when audit resources cannot be adapted in the short term.

Erard & Feinstein proposed a solution method leading to two differential equations from which $x(y)$ and $p(x)$ can be calculated. We now claim that the E&F solution method is incorrect. We use instead a simple simulation algorithm to determine the functions $x(y)$ and $p(x)$ which approximate the optimal functions.

The computer code AUDSIM uses a class of $p(x)$ functions with two or three parameters. For each set of parameters, i.e. the $p(x)$ function, the auditee behaviour is simulated, providing a function $x(y)$. The parameters in the $p(x)$ function are varied until the objective function is minimized,

This optimization problem may have multiple local optima, which complicates the solution method as multiple minima have to be evaluated and compared. Examples of multiple minima are shown in Section 5.2.

5 Results

5.1 Base case: Group B unsegmented

Today there is no sanction in Sweden for benefit fraud corresponding to the tax surcharge. The sanction factor *a* is therefore set to zero in the Base case.

The distribution for groups ABC in Figure 2 has been used in the determination of optimal audit strategy together with share of error-free persons *Q = 1 – 0.112 = 0.888* according to Section 4.2 above.

Figure 4 shows optimal relative audit frequency functions for different audit unit cost levels. For unit costs exceeding SEK 1 000, the normal type of strategy is optimal, whereas an alternative strategy is optimal for audit unit costs below SEK 1 000. This alternative strategy implies 100 % audit of all persons with declared claims > 0. In this example, there is a continuous transition from the normal strategy to the alternative strategy, as illustrated by the curve for audit unit cost SEK 1 010 which has the normal shape although close to the alternative strategy.

In the normal strategy, no audits shall be made below a certain level of declared claims. This cut-off level depends on the audit density and varies between SEK 70 at audit unit cost SEK 1 010 and SEK 5 200 at audit cost SEK 3 000.

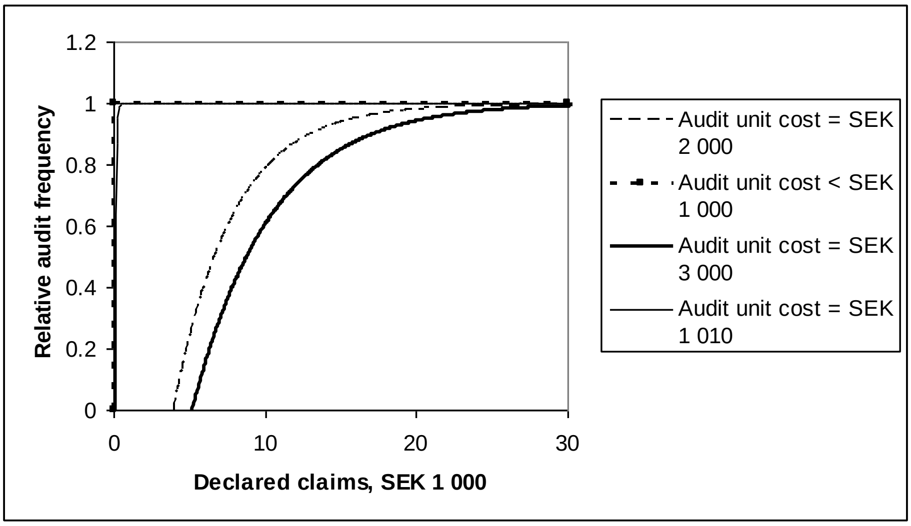


Figure 4. Optimal audit frequency functions for group B with varying audit unit cost

Fraudsters with true claims below the cut-off level make claims close to the cut-off level.

In the computation of optimal audit frequency, the optimum fraud amount is also calculated, i.e. the difference between declared claims and true claims. Optimal fraud amount for group B is shown in Figure 5 as a function of true claims for varying audit unit costs.

For low audit unit costs, the alternative strategy uses 100 % audits for all positive claims. In this case, all fraud amount functions are optimal.

The fact that the fraud amount functions in Figure 5 are piece-wise linear is a consequence of the selected class of audit frequency functions. The truly optimal fraud functions are most probably non-linear.

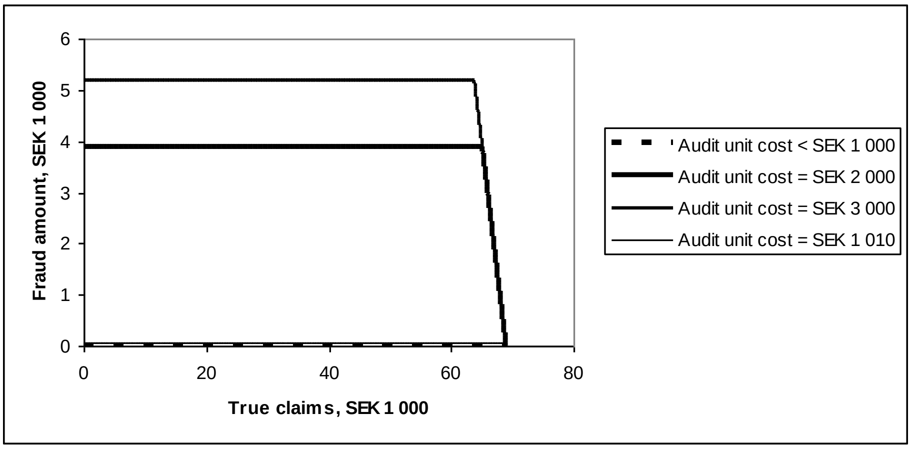


Figure 5. Optimal fraud amount as a function of true claims for group B at varying audit densities

Figure 6 shows the average fraud cost per person with positive claims as a function of audit density. This is the cost of remaining undiscovered fraud if the fraudsters act rationally and the auditor uses the optimal strategies shown in Figure 4.

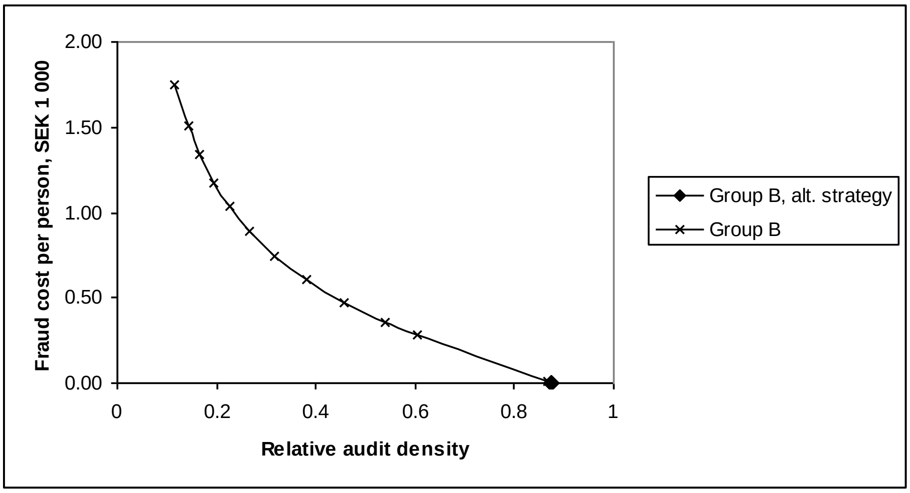


Figure 6. Average fraud cost per person with positive claims as a function of audit density for group B

Figure 7 shows the fraud cost and the total cost, i.e. the sum of fraud cost and audit cost, per person with positive claims as a function of audit unit cost. For unit costs less than SEK 1 000, the total cost curve is linear since the fraud cost is zero in the alternative strategy.

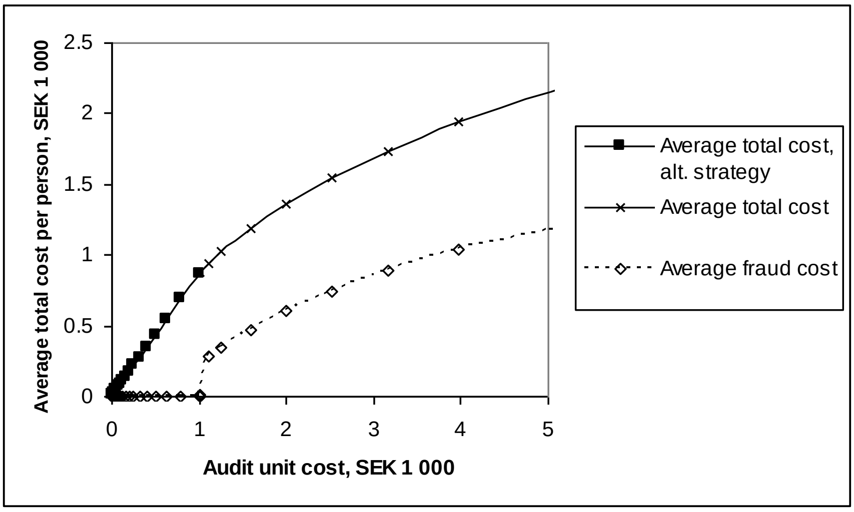


Figure 7. Average total cost and average fraud cost per person as a function of audit unit cost for group B

It may seem surprising that the fraud cost function seems to be discontinuous at an audit unit cost around SEK 1 000, with no corresponding discontinuity in the total cost. The explanation is a compensating near-discontinuity in the audit density, as shown in Figure 8.

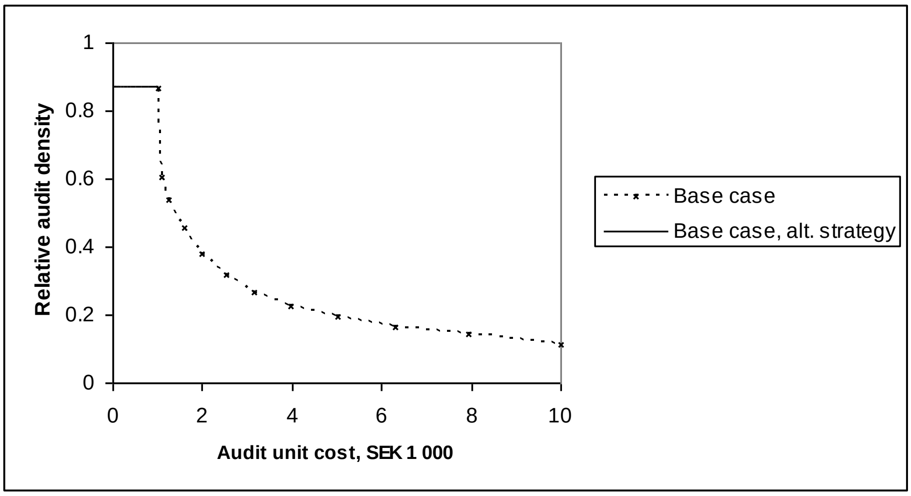
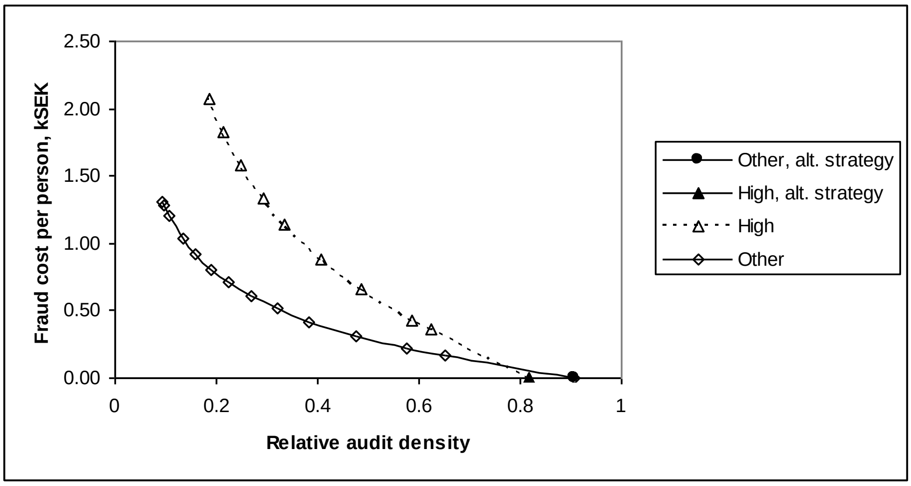


Figure 8. Relative audit density as a function of audit unit cost, group B

5.2 Segmentation of group B

Group B is segmented according to Section 3 in the segments *High risk* and *Other*. If the audit unit cost is externally determined, the optimal audit strategy is determined for each segment using the audit unit cost, the fraction of honest auditees and the distribution of true claims. In case the total number of audits is determined externally, the audit unit cost is replaced by a Lagrange multiplier λ which is varied until the total number of audits in all segments matches the audit constraint.



Figur 9. Fraud cost per person with positive claims as a function of audit density for the segments *High risk* and *Other*

Figure 9 shows the average fraud cost per person with positive claims for the two segments. As in the Base case, the alternative strategy is optimal for high audit densities.

Figure 10 shows the fraud cost and the total cost per person with positive claims as a function of audit unit cost. As long as the alternative strategy is optimal, i.e. up to audit unit cost = SEK 700 for the *Other* segment and up to SEK 1 800 for the *High risk* segment, the total cost function is linear in Figure 10.

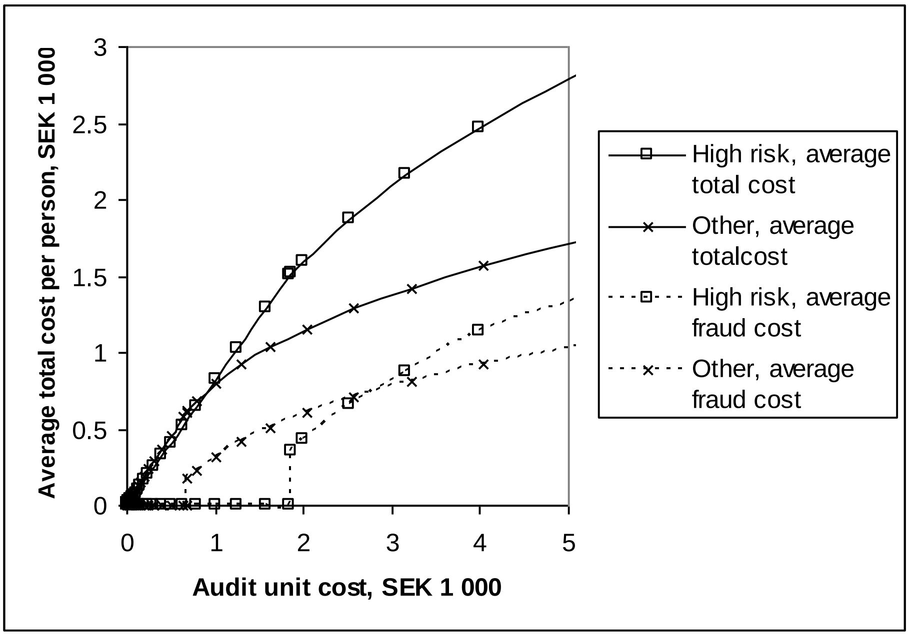


Figure 10. Total cost per person with positive claims as a function of audit unit cost for the segments *High risk* and *Other*

Figure 11 reveals that the audit density is discontinuous between the ordinary and the alternative strategy for both segments, in contrast to the Base case in Figure 8. In case an auditor wishes to use an audit density in the missing interval, this can be accomplished with a linear combination of the end-point strategies.

For both segments, we have a case with two local minima. Therefore, the total cost for the normal and the alternative strategy have to be compared after which the strategy with the lowest cost is selected.

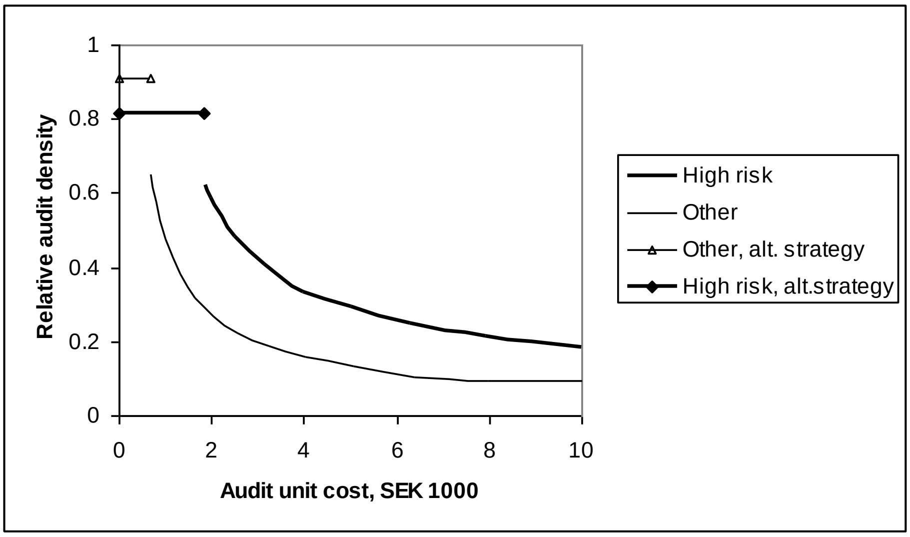


Figure 11. Relative audit density as a function of audit unit cost for the two segments

5.3 Combination of the two segments

In Table 6, the average total cost is calculated for the segmented case when the two segments are combined. The total cost function is compared with the Base case function in Figure 12. For low audit unit costs where the alternative strategy is optimal for both segments, segmentation gives no cost reduction. For higher audit unit costs, a cost reduction up to 8% is achieved.

Table 6. Calculation of combined average total cost

Relative audit unit cost, SEK 1 000	*Average total cost per person, SEK 1 000*		
	High risk, 626 persons	*Other, 1 083 persons*	*Combined, 1 709 persons*
0.05	0.041	0.051	0.048
0.1	0.081	0.098	0.092
0.316	0.259	0.295	0.281
1.0	0.817	0.792	0.801
3.16	2.16	1.41	1.69
10	3.92	2.23	2.85

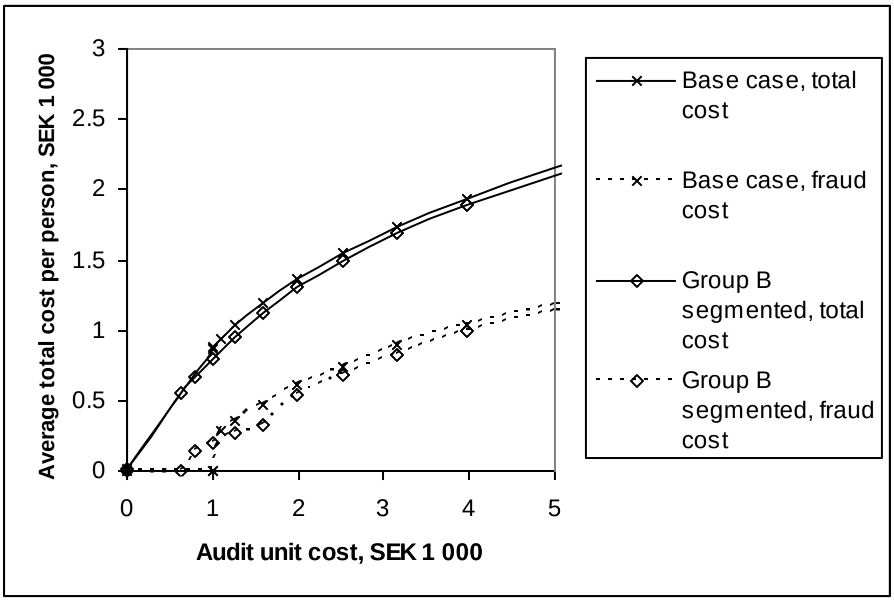


Figure 12. Average cost per person with positive claims as a function of audit unit cost for Group B segmented compared to the Base case

In Table 7, the relative audit density and the average fraud cost is calculated for the segmented case when the two segments are combined. The relation between audit density and fraud cost is compared to the Base case in Figure 13. The cost reduction obtained with segmentation is small for low audit densities but exceeds 30 % for relative audit densities above 0.6.

Table 7. Calculation of combined audit density and fraud cost

Relative audit unit cost, SEK 1 000	*Relative audit density*			*Average fraud cost per person, SEK 1 000*		
	High risk	*Other*	*Combined*	*High risk*	*Other*	*Combined*
0.05	0.817	0.908	0.875	0	0	0
0.1	0.817	0.908	0.875	0	0	0
0.316	0.817	0.908	0.875	0	0	0
1.0	0.817	0.476	0.601	0	0.316	0.200
3.16	0.407	0.191	0.270	0.875	0.806	0.831
10	0.185	0.094	0.127	2.074	1.306	1.587

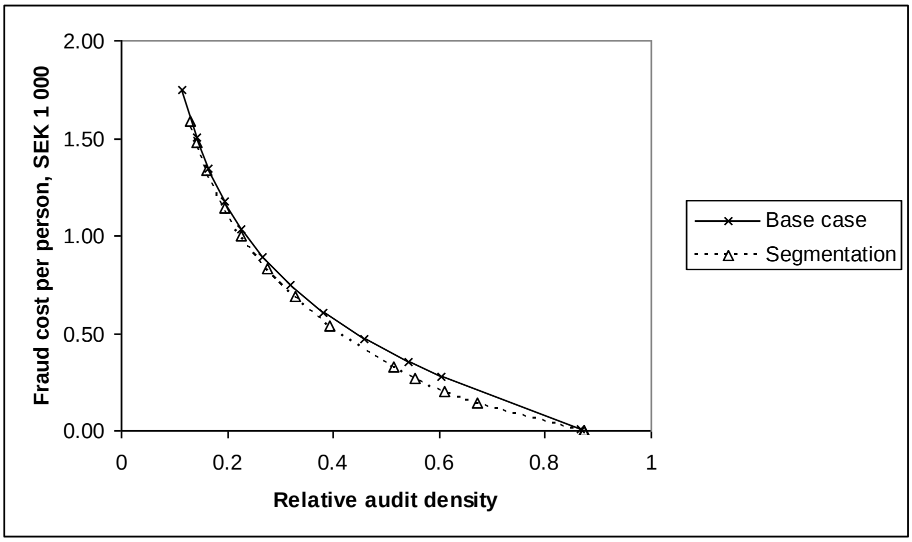


Figure 13. Fraud cost per person with positive claims as a function of audit density for Group B segmented compared to the Base case

According to Table 7. the *High risk* segment should be audited with about twice the audit density compared to the *Other* segment for relative audit costs exceeding SEK 1 000.

In Figure 14, the optimal audit frequency functions are shown for audit unit cost = SEK 1 000.

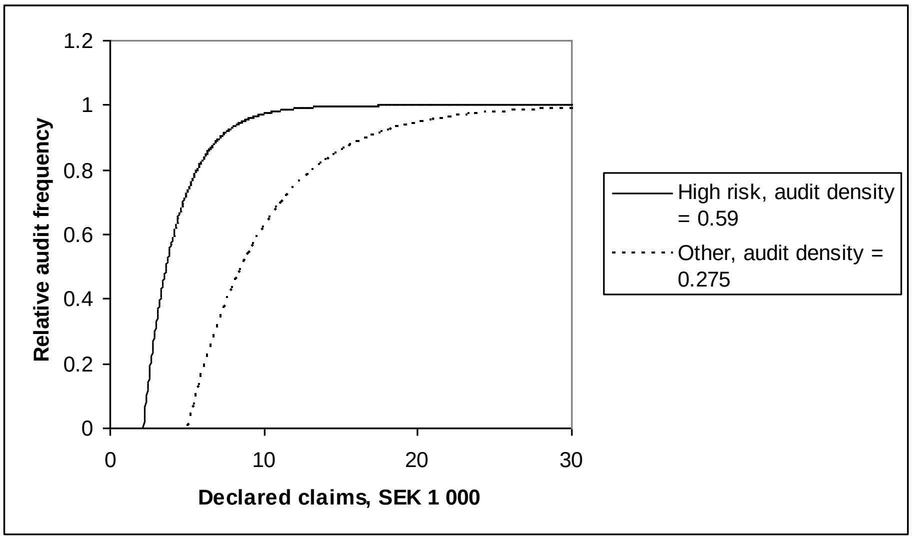


Figure 14. Optimal audit frequency functions for the two segments at audit unit cost = SEK 2 000

Figure 15 shows the optimal relative audit density as a function of audit unit cost for the combined segments compared to the Base case. The irregularity in the curve for the segmentation case depends on the jump between ordinary and alternative strategy for the *High risk* segment shown in Figure 11.

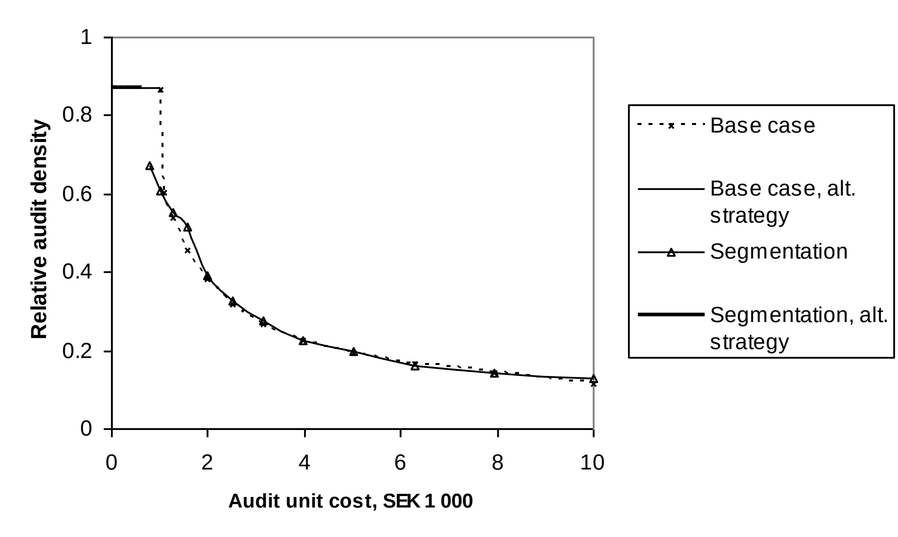


Figure 15. Relative audit density as a function of audit unit cost for the combined segments and the Base case

5.4 The effect of sanctions

Today, there is no sanction against social benefit fraud in Sweden corresponding to the tax surcharge**.** Such a sanction as been proposed in 2011 in a government study, with a surcharge amounting to 20% of the faulty claim, combined with a lower and an upper limit.

Even without such a sanction, it is possible that a fraudster experiences the risk of discovery and prosecution for benefit fraud as a social sanction with a similar effect as a surcharge. It is therefore of interest to study how the optimal audit strategies are affected by a surcharge proportional to the size of the error.

The effect of a surcharge (*sanction factor*) of 25 % of the discovered error in the Base case is studied below. It has no other effect on the audit strategies than that the relative audit density/frequency no longer coincides with absolute audit densities/frequencies. Instead, the conversion factor 1/(1+a) = 0,8 is applied, i.e. that the fraud cost in Figure 6 is reduced to zero at a 20 % lower audit density and that the audit frequency functions in Figure 4 instead has a maximum value of 0.8 as in Figure 16.

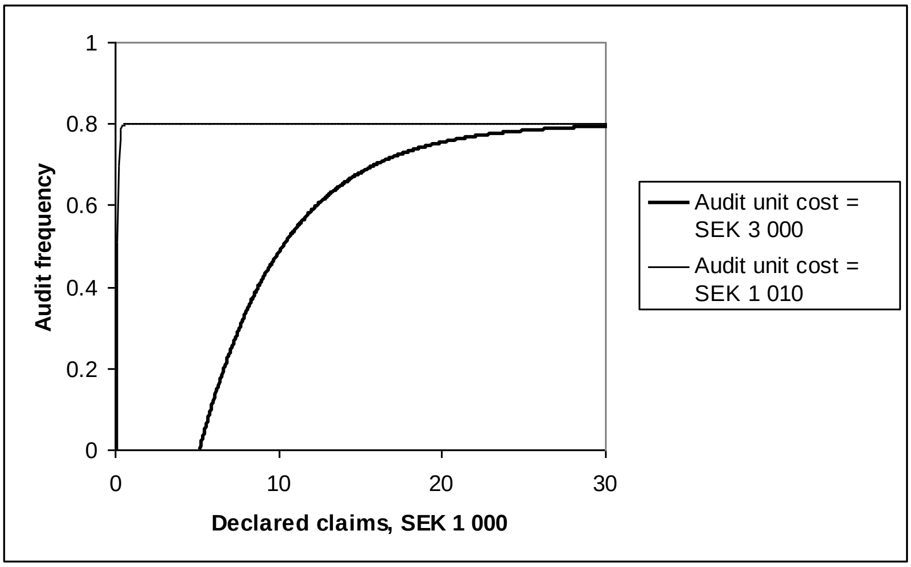


Figure 16. Optimal audit frequency functions for group B with varying audit unit cost, sanction factor 25 %

In Figure 17, the average fraud cost is shown as a function of audit density with a 25 % sanction factor, compared to the cost without sanctions from Figure 6. The curve is compressed to the left with a factor 0.8. This figure is relevant when the audit density, i.e. the number of audits, is determined exogenously.

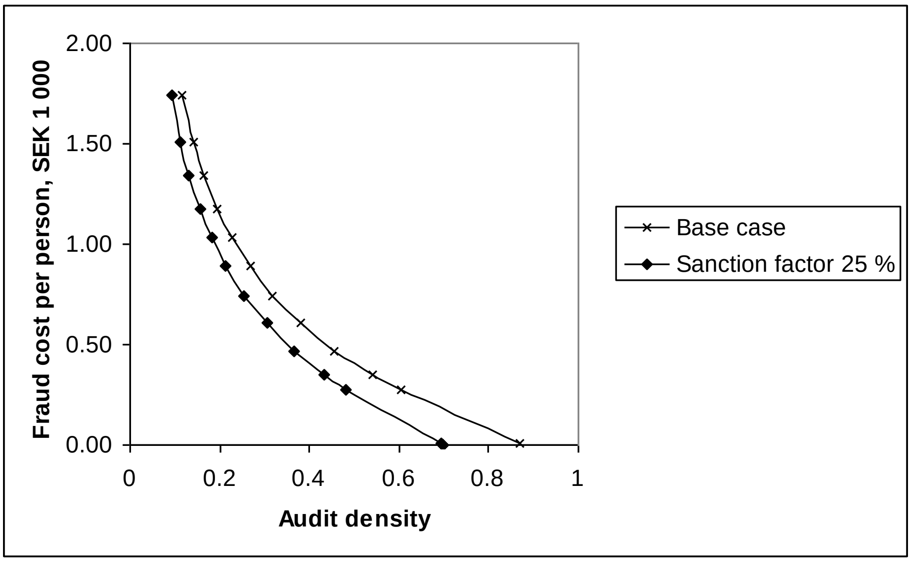


Figure 17. Average fraud cost per person with positive claims as a function of audit density for group B, sanction factor 25 %

In Figure 18, total cost and fraud cost as a function of the audit unit cost with a 25 % sanction factor is compared to the corresponding data from Figure 7. The total cost is reduced with 20 % for audit unit costs below SEK 1 000 where the alternative strategy is optimal. For higher audit unit costs, the reduction falls to about 10 %. It is thus apparent that an introduction of a sanction fee will lead to considerable savings. This figure is relevant when the audit unit cost is determined exogenously.

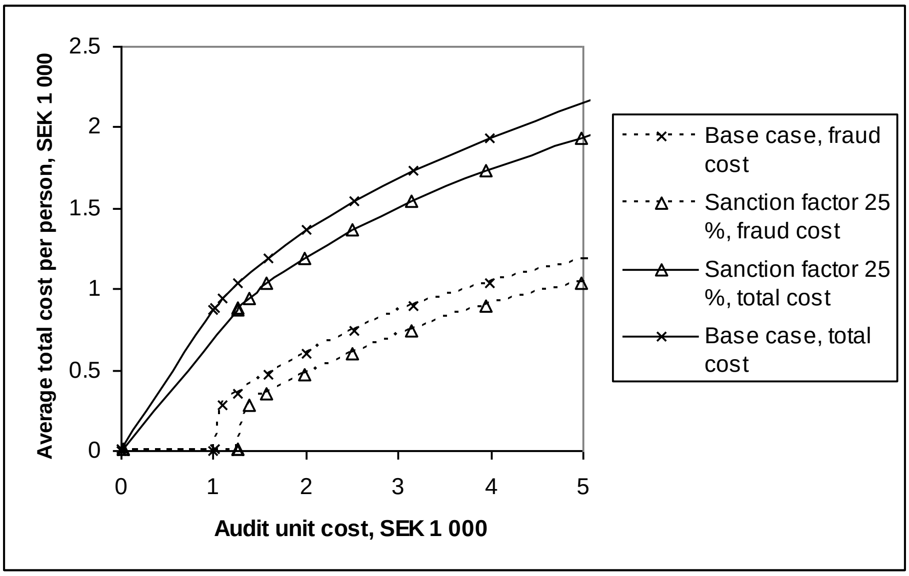


Figure 18. Average total cost and average fraud cost per person as a function of audit unit cost with a 25 % sanction factor, compared to the Base case

Figure 19 shows the optimal audit density as a function of audit unit cost for the Sanction case compared to the Base case. The curve for the Sanction case is a transformation of the Base case curve, with a compression with 20 % along the y axis and an enlargement with 25 % along the x axis.

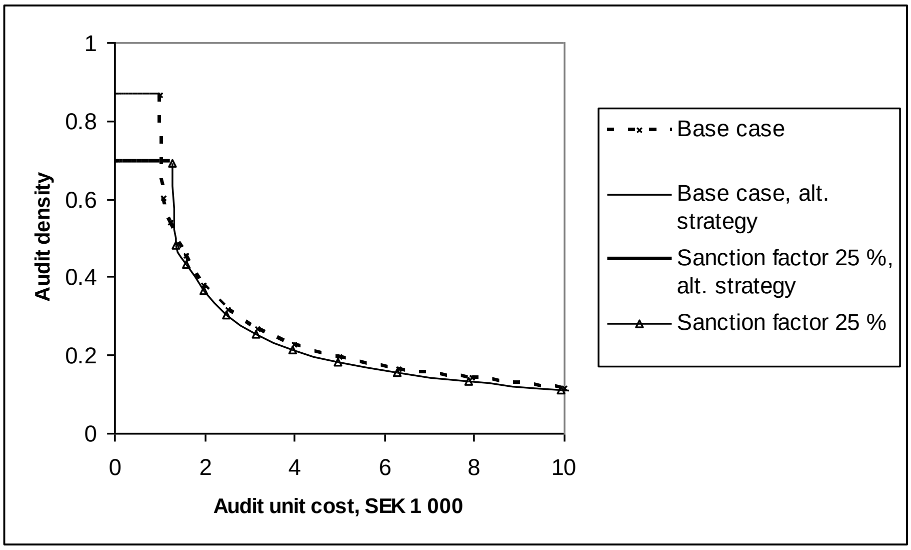


Figure 19. Relative audit density as a function of audit unit cost with a 25 % sanction factor compared to the Base case

6 Results with a mix of fraud and involuntary errors

6.1 Background

FK made in 2007 a study of TFP on 779 persons in the Gothenburg and Jönköping regions (Försäkringskassan 2007), where telephone audits were made with employers and schools/pre-schools. For 14.8 % of the persons, an error was recorded, i.e. involuntary error or fraud. After investigation, suspicion of fraud remained with 7.1 % of the population. This supports the FK opinion that fraud and involuntary errors are approximately of the same magnitude.

6.2 Model

The errors which are discovered in an audit are assumed to partly depend on fraud and partly on involuntary errors, in the proportions A and $1-A$, respectively. As in the E&F model, the amount of fraud is assumed to be affected by the audit strategy, such that the fraudster selects the fraud amount which maximizes his/her expected utility. The involuntary errors are assumed to be determined exogenously and are not affected by the audit strategy. The optimal audit strategy is affected by the involuntary errors since the objective of the auditor is to discover both fraud and involuntary errors.

A simple model for involuntary errors is to assume that a fraction of those who are not fraudulent make involuntary errors proportional to their total claims. A consequence of such a model is that those with true claims equal to zero would not make any involuntary errors. As this is contradicted by the empirical data where about 50 % of the erring persons have zero true claims, this model must be discarded.

Another simple model is to assume that a fraction of those who are not fraudulent make an error of a fixed amount. This model has been included in the code AUDSIM where the fraction of erring non-fraudsters and the size of the error are two additional parameters.

A cost term is added in the model

$$C_2 = -Q_a E Q_e \int (1-(1+a)p(u(y)))f(y)dy$$

where E is the constant error amount for the erring fraction Q_e of the non-fraudsters.

u and y are declared and true claims for the erring non-fraudsters with the relation $u = y+E$

$Q_a = Q + (1\text{-}Q)(1\text{-} A)$ is the fraction of non-fraudsters in the total population whereas Q as before is the fraction of error-free persons, determined from the empirical data.

$Q_e = (1\text{-}Q)A/Q_a$ is the fraction of non-fraudsters who make involuntary errors.

We assume arbitrarily that the error amount is shared between fraudsters and persons with involuntary errors in proportion to their number, i.e. that the fraudsters are responsible for the fraction A of the total error amount. The error amount E is then equal to the average error per erring person, calculated in Table 1.

6.3 Results

The fraction of erring persons amounts to 11.2 % in group B. If we assume that 50% of the errors are involuntary, the fraction of fraudsters would be 5.6 % of the population, i.e. Q_a = 0.944. The fraction of non-fraudsters making involuntary errors is Q_e = 0.056/0.944 = 0.0593. The error amount E is SEK 5 759 according to Table 1.

Optimal audit frequency functions at 50 % involuntary errors are shown in Figure 20. For audit unit costs less than SEK 800, the alternative strategy is optimal.

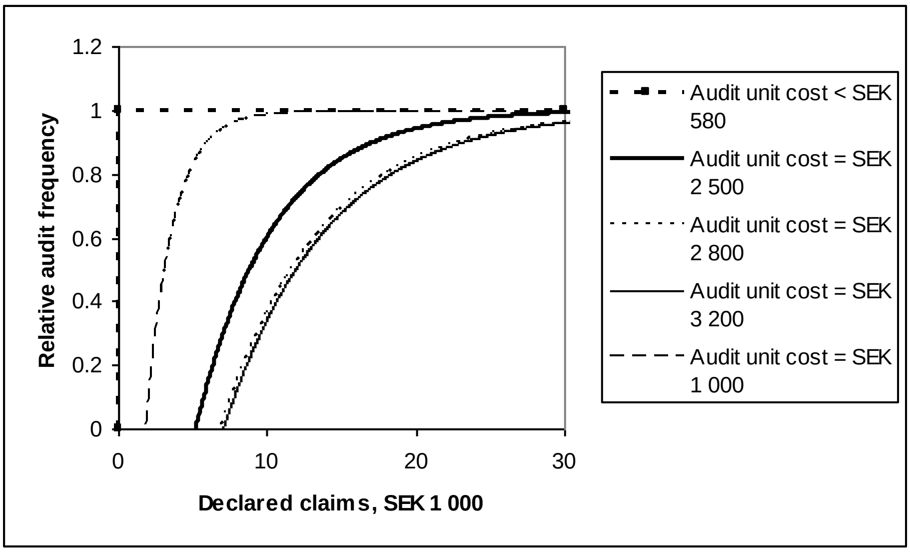


Figure 20. Optimal audit frequency functions with 50 % involuntary errors and varying audit unit cost

Figure 21 shows the optimal fraud amount as a function of true claims for the five cases in Figure 20. In Figures 20 and 21, it is apparent that the functions are quite similar for audit costs above SEK 2 800. We have no simple explanation to this phenomenon.

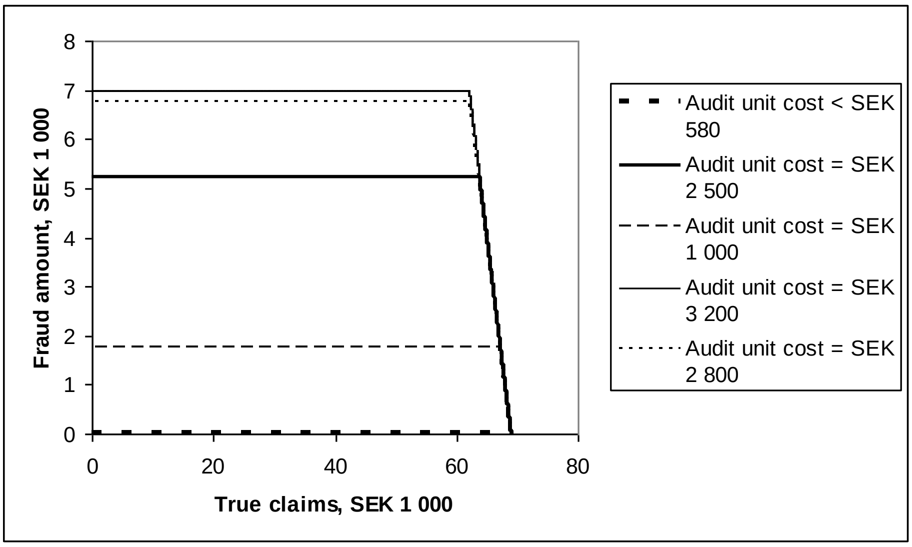


Figure 21. Optimal fraud amount as a function of declared claims with 50 % involuntary errors and varying audit unit cost

Figure 22 shows the error cost, i.e. the sum of fraud cost and the cost of involuntary errors, as a function of relative audit density at 50 % involuntary errors. The fraud cost curve for the Base case (Figure 6) is included for comparison.

Compared to the Base case, the error cost is lower at low audit density and higher at medium high audit density. This is natural since the amount of fraud adapts to the audit density and is thus extra high for low audit densities, whereas the involuntary errors are independent of audit density such that the cost of remaining errors varies linearly with audit density. At the high end, the two curves coincide since the alternative strategy is optimal for both cases.

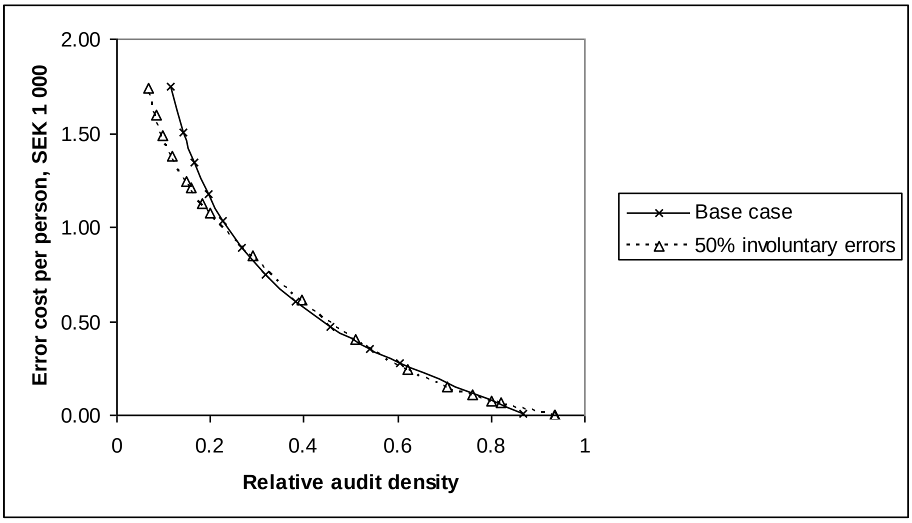


Figure 22. Average error cost per person with positive claims as a function of audit density with 50 % involuntary errors, compared to the Base case

Figure 23 shows the average cost per person with positive claims as a function of the audit unit cost. The case with 50 % involuntary errors is compared to the Base case (Figure 7). As expected, the total cost is higher in the Base case for high audit unit costs, corresponding to low audit densities whereas the opposite is true for medium high audit costs.

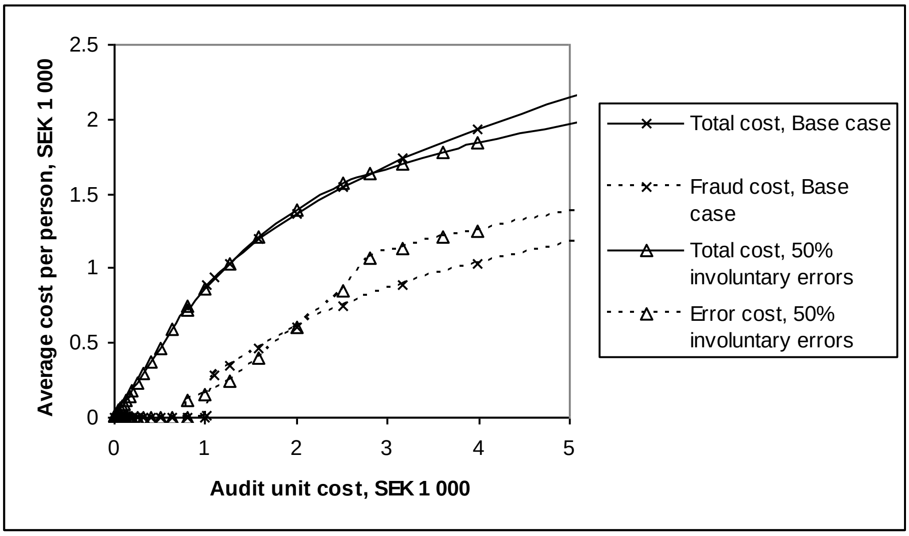


Figure 23. Average total cost and average fraud cost per person as a function of audit unit cost with 50 % involuntary errors, compared to the Base case

For low audit unit costs, up to SEK 580 where the alternative strategy is optimal, the total cost is about 7 % higher than in the Base case. This is due to the fact that a larger number of persons must be audited in the alternative strategy, since the persons who make involuntary errors continue to do so even if the audit frequency is 100 % as is the case in the alternative strategy, whereas the fraudsters abstain from fraud in the alternative strategy.

Figure 24 shows optimal relative audit density as a function of the audit unit cost. The case with 50 % involuntary errors is compared to the Base case (Figure 8). The curve for involuntary errors is lower for high audit unit costs and higher for low audit costs. This is natural since in a case with 100 % involuntary errors, no audits would be made for audit costs exceeding the average error amount whereas 100 % audits would be made for lower audit costs.

As explained above, the optimal audit density is higher compared to the base case in the interval where the alternative strategy is optimal, i.e. up to SEK 580.

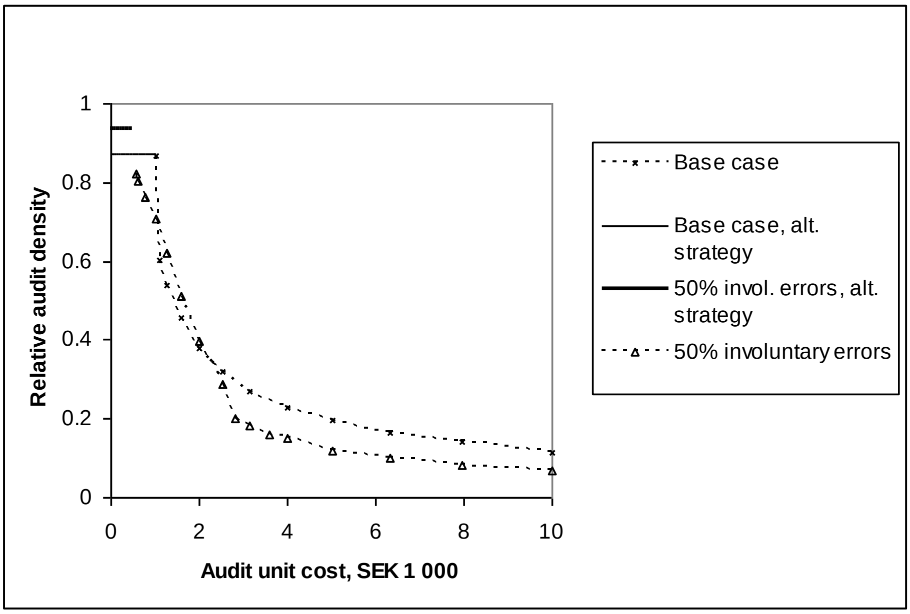


Figure 24. Relative audit density as a function of audit unit cost with 50 % involuntary errors, compared to the Base case

7 The cost of audits

7.1 General

The estimates of audit costs for schools, pre-schools and employers are uncertain since they are based on a few interviews with representatives for each category.

The estimates concerning FK should be more accurate since they were reported by FK audit management. On the other hand, quite different estimates were received from FK during the course of the project.

Wherever possible, the cost estimates have been made in the 2006 cost level since the claim amounts refer to the time period when the IFAU study was carried out. In case the results shall be used for future audit design, the cost estimates have to be updated, of course.

7.2 Audit cost within the Swedish Social Insurance Agency (FK)

Time requirement for telephone audits

In the spring of 2010, the audit management within FK estimated the time required for a telephone audit to 20 minutes. The staff cost per hour was around SEK 300 according to FK.

A higher cost is obtained if the total management cost in FK, amounting to SEK 9 225 million, is divided by the number of equivalent full time employees, 14 038 persons. Given 1 750 work hours per annum, the hourly cost should be around SEK 375.

In October, 2010, additional questions were brought up with FK. It was then reported that the time requirement was now reduced to 5 minutes for a request regarding one claim. If a request is made concerning several claims by one person, the additional time requirement is estimated to be 2 minutes per claim. It is assumed here that the estimates latest received from FK are the most accurate.

The average number of claims per person during the total period is 5.9 according to Table 1. The time requirement for audit would then be five minutes for the first claim plus 4.9 times two minutes for the remainder, totally 14.8 minutes, rounded to 15 minutes.

Audits against employers versus audits against schools/pre-schools

In an audit study made by FK (FÖRSÄKRINGSKASSAN 2007), audit calls were made to both employers and schools/pre-schools. In some instances, no answer was obtained. For those persons where audits were successful for both employers and school/pre-school, results were obtained according to Table 8.

Table 8. Error shares in the FK audit study

	Total error share, %	*Fraud share, %*
Errors at employer and school/pre-school	2.1	1.4
Error at employer only	7.0	3.6
Error at school/pre-school only	2.8	0.6

According to the FK study, the most common fraud is to claim TFP while still at work but letting the child be absent from school/pre-school.

As a vast majority of errors (82 % of total errors, 75 % of fraud) occurs at one place only, it is obvious that audits should be carried out at both employer and school/pre-school. It is therefore assumed below that the audit of one person consists of two telephone calls.

Audit cost

Given an hourly cost of SEK 300-375, the cost for 2x15 minutes amounts to SEK 150-188.

7.3 Audit cost in schools/pre-schools

Time requirement

Five interviews were conducted with school/pre-school staff. The conclusions are that information is available at least six months back in manual or computerized journals, and that the time requirement is 5-10 minutes for one claim and 2-4 minutes per additional claim.

Audit cost

According to Statistics Sweden (SCB), the average monthly salary of city employees was SEK 21 284 in 2006. This is judged to be a good estimate for teachers in comprehensive school and pre-school staff. Assuming 50 % indirect employee costs and 1 750 work hours per annum, the hourly cost is around SEK 220.

For the audit of claims during a period of 8.5 months, the time requirement is calculated to 15-30 minutes in the same way as in Section 7.2. The cost is SEK 55-110 per audited person.

7.4 Audit cost for employers

Three interviews were conducted, at one large, one medium-sized and one small company.

The time requirement for audit of one claim where the information is computerized is estimated to be five minutes, whereas it is 15 minutes if the information has to be obtained from manual files. Based on the calculation in Section 7.2, the time required for claims during an 8.5 month period is 15-45 minutes. The hourly cost is estimated to be the same as for FK, i.e. SEK 300-375. The cost interval for employer audits thus is SEK 75-281.

7.5 Total audit unit cost for telephone audits

The audits for an 8.5 month period generate costs according to Table 9.

Table 9. Audit unit costs

Organization	*Unit cost, SEK*
FK	150-188
School/pre-school	55-110
Employers	75-281
Total	280-580

8 Total cost and optimal audit density

General

As the maximum audit unit cost in Table 9 is less than SEK 600, all optimum solutions belong to the alternative strategy, with 100 % relative audit frequency for all persons declaring positive claims. This means that the cost of remaining errors/fraud is zero in all cases, such that the only cost is the cost of auditing.

With such a strategy, those persons who had zero true claims and made fraudulent declared claims will abstain from fraud. The audit density, measured as the ratio between persons audited and persons with positive claims in the sample, will therefore be less than 100 %.

Optimal audit densities and average total costs are presented in Table 10 below.

Base case

Average total cost for various audit unit costs is obtained from Figure 7. Optimum relative audit density is obtained from Figure 8.

Segmentation

Average total cost for various audit unit costs is obtained from Figure 12. Optimum relative audit density is obtained from Figure 13. As the optimum is of the alternative type, no cost savings are obtained.

Sanctions

With a 25 % sanction factor, i.e. a surcharge amounting to 25 % of the error amount, the average total cost is reduced considerably according to Figure 17. The relative audit density also falls considerably according to Figure 18.

Involuntary errors

With a 50/50 split between fraud and involuntary errors, the average total cost per person is about 7 % higher compared to the Base case for audit unit costs below SEK 580 (Figure 23). Also the optimal audit density is about 7 % higher compared to the Base case (Figure 24).

The Combination case

In the analysis above, it has been stated that segmentation leads to slightly lower costs, and that the introduction of a sanction fee leads to considerably lower costs. It has also been stated that a mix of fraud and involuntary errors is a realistic assumption. Such a mix sometimes leads to higher and sometimes to lower costs.

ISF has expressed the view that the most realistic combination of the cases studied above is a case with segmentation, sanctions (25 % sanction factor) and a mix of fraud and involuntary errors (50/50). This case is denoted the Combination case.

Summary of results

The main results with the different cases studied in Sections 5 and 6 and the audit unit costs estimated in Section 7 are summarized in Table 10.

Table 10. Summary of results concerning optimal audit density and average total cost per person with positive claims

	Internal costs only SEK 150-188		*Internal plus external costs SEK 280-580*	
	Optimal audit density	Total cost per person	Optimal audit density	Total cost per person
Base case	87 %	SEK 130-165	87 %	SEK 245-505
Segmentation	87 %	SEK 130-165	87 %	SEK 245-505
25 % sanction factor	70 %	SEK 105-130	70 %	SEK 195-405
50 % involuntary errors	93 %	SEK 140-175	93 %	SEK 260-540
Combination case	75 %	SEK 115-140	75 %	SEK 210-435

With internal costs only, the average total cost per person is SEK 130-165 in the Base case according to Table 10. The annual cost per person becomes 12/8.5 times larger, i.e. SEK 184-233. As 683 000 persons claimed TFP in the year 2006, the annual cost would be SEK 125-160 million.

A corresponding calculation including audit costs for schools/pre-schools and employers renders an annual cost of SEK 235-485 million.

The total TFP payments amounted to about SEK 3 900 million in 2006. The annual cost for audits and remaining errors would in the Base case vary between 3 and 12 % of the total payments, which seems to be a realistic share.

The annual cost for an optimal audit system is as shown above in the range SEK 125-485 million. This can be compared with the annual cost of errors calculated in Section 2, i.e. SEK 660 million, which does not include the actual audit cost incurred in 2006. It is obvious that large savings are possible with an improved audit system.

A reporting system for schools/pre-schools has been introduced since 2006, which results in almost 100 % verification of absence from school/pre-schools. This is in line with the results of this study that a high level of audits or verification is desirable, subject that the costs for the reporting system are comparable with the audit costs. No corresponding action has been taken concerning employers, however, implying a gross imbalance which may encourage certain types of fraudulent behaviour.

Sensitivity analysis

Figure 25 shows the total cost in the Combination case if a non-optimal audit density is chosen instead of the optimal audit density level. For the audit unit cost SEK 400, the optimum is at the extreme right end, and it is obviously quite costly to deviate from the optimum audit density. For the audit unit cost SEK 1 500, however, the optimum is in the interior of the audit density interval, making the optimum extremely flat and the cost for deviating from the optimum quite small.

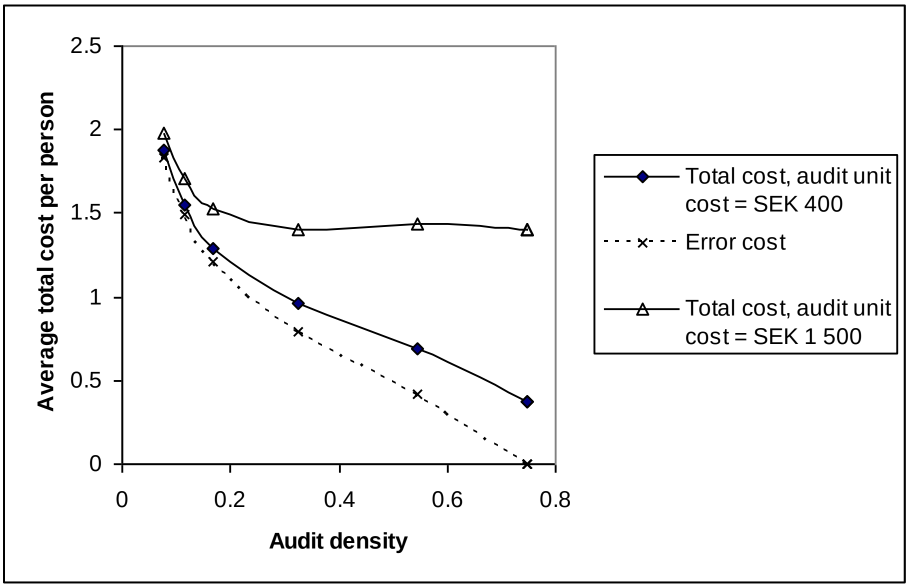


Figure 25. Average cost per person with positive claims as a function of audit density in the Combination case

9 Discussion

9.1 Goal satisfaction

The analysis above shows that it is possible to apply the Erard & Feinstein model on social benefit fraud.

Most of the analyses concern the case when all errors are caused by fraud. The level of fraud is assumed to be affected by information regarding the audit strategy, i. e. that persons with large TFP claims will be audited more frequently. However, audits made by FK indicate that about 50% of the errors are involuntary. A model which includes involuntary errors has been developed as part of this study.

The analysis has led to quantitative results which make it possible to determine a near-optimal audit frequency function and a near-optimal audit density for a given audit unit cost. It is also possible to compare the results of this study with other existing or proposed audit systems, for example the reporting system used for verification of TFP claims towards schools/pre-schools.

The project has not aimed at a validation of the model used, i.e. to study how persons claiming TFP would react to audit strategy information. Such validation would require additional empirical data.

9.2 Data base

The data base from the IFAU study has been a useful basis for the analysis. However, it would have been a considerable improvement if FK had also audited the control group unaffected by information (group D).

The population audited has been augmented with persons with zero true claims and zero declared claims in order to satisfy the model requirement that persons with zero true claims shall have the same error propensity as persons with positive true claims. This should be regarded as a technical correction. It is quite probable that persons with zero true claims have a different error propensity compared to persons with positive true claims.

The number of erring persons with zero true claims has not been used as an indicator for the segmentation. This may have reduced the quality of the segmentation.

Audit data from all three groups have been used for the segmentation in order to create a larger data base. The groups A and C have a larger average true claims amount compared to group B. In order to give the claims distribution its correct mean value, the claims distribution should be compressed with the ratio between average true claims for group B and average true claims for groups ABC, i.e. the factor 1410/1454 = 0.97. This correction has not been made in this study.

It is a matter of discussion whether data from the reference period should be used or not. One reason for using such data is that a larger data base is obtained, which hopefully results in a more correct true claims distribution. On the other hand, the data from the reference period contains declared claim data only, and no error data obtained from audits. If data from the reference period are used, it is necessary to supplement them with assumption regarding errors. It is natural to assume that each individual has the same relative error amount during the reference period as during the audit period. This assumption is not plausible for groups A and C, however, since they have been subjected to a warning letter. In the study, this assumption has been made also for groups A and C in order to create a larger data base for the full period. The effect on the true claims distribution is estimated to be small.

The method for recalculation of the fraction of erring persons with the factor 1.411 can be discussed (Table 4 in Section 3). Such a method for adjustment could theoretically lead to a fraction of erring persons exceeding 100 %. Given the low levels used here, the method should be acceptable.

9.3 Model validity

The separation in rational fraudsters and inherently honest persons is of course a simplification. It is probable that persons exist between those two extremes who decrease their amount of fraud or abstain from fraud because of the social sanction imposed by discovery and possible prosecution. Such behaviour can probably be included in the model through the introduction of subgroups with varying social costs in addition to the monetary cost of loss of benefit and possible monetary sanction. Estimating such a model empirically of course represents a challenge.

According to Table 1, groups A and C have a smaller share of erring persons than group B. This contradicts the model, where it is assumed that all the groups are drawn from the same population with a given fraction of honest persons. In the E&F model, only the fraud amount varies with the perceived audit density. A possible explanation is that all groups have the same fraction of inherently honest persons. In addition, some individuals may abstain from fraud because of the perceived higher audit density caused by the warning letter.

With such an explanation, it is possible that the inherently honest fraction in group B should be lower, i.e. that persons exist in group B who abstain from fraud because of the perceived risk of discovery.

Fraud amount in model and reality

In the application of the E&F model, the error amounts have not been used in order to calibrate the model. According to the model, the fraction of fraudulent persons is exogenously given, whereas the fraud amounts depend on the audit frequency. Consequently, all persons with the same true claims will have the same fraud amount, which definitely is not the case in the empirical data. The model assumes that the fraudsters have full information regarding the audit frequency function, which of course is not the case in reality. Each individual has a personal perception of the audit frequency and its dependence on the declared claims, which leads to large variations in the fraud amounts in the empirical data.

Audit quality

It is assumed in the model that all errors are discovered in an audit. This is not the case in real life. For instance, it is very difficult to discover "care of healthy child", which would require home visits by medically educated staff.

Although normal audit methods will not discover all types of errors, the E&F model is useful since it can be applied on the auditable errors which most probably are the majority. Theoretically, improved audit methods for auditable fraud could lead to a shift towards fraud types which cannot be audited. An example is the present reporting system for schools/pre-schools where the 100 % verification could lead to increased fraud on the employer side where the audit density is low. The parent can claim TFP but continue to work while the child is alone at home, accompanies the parent at work or is taken care of by another person. An increasing number of employees also have the possibility of working at home using PC's, internet etc.

The risk of "negative fraud", i.e. underutilization of benefits

The IFAU paper resulted in considerably higher error levels compared to FK estimates. A possible explanation which has been brought forward by critics of the IFAU study is that some persons were so scared by the warning letter that they abstained from claims in order to avoid the risk of making errors and be pointed out as fraudsters. This effect is most probable for persons with insufficient knowledge of the regulations.

The same risk exists if a high audit frequency is announced for persons with large claims. This risk is higher for less educated persons with many children or children with impaired health. The use of the audit strategies recommended in this paper should therefore be combined with information and education activities.

Linear or concave utility function

The fraudulent persons are assumed in this paper to have a linear utility function. This assumption seems reasonable since the errors are generally small, with an average below SEK 6 000**.** For those with the largest claims, of the order of SEK 50 000, the assumption may be doubtful.

Erard & Feinstein extended their model and solution method to a concave utility function. This can easily be implemented also in the AUDSIM computer code.

9.4 Future research

Possible extensions of the TFP project

- The allocation of audits between employers and school/pre-schools can be optimized, taking into account differences in audit unit costs.
- The economics of the audit strategies described in this paper can be compared to the economics of the present reporting system for schools/pre-schools.
- The segment *Stockholm county* has a surprisingly high share of persons with errors. This segment can be split up into smaller sub-segments in order to obtain an increased understanding of the risk factors, if possible. Such a study can easily be carried out since the data base contains information regarding municipality and parish within municipality.

Model development

- The model can be extended to include a concave utility function.
- The model can be extended to include groups of persons with behaviour between the rational fraudsters and the inherently honest persons, for instance with various levels of social sanction costs attached to audits and discovery.

Other possible ISF projects

- Another benefit of larger economic significance can be studied. The heaviest benefits within FK are Sickness cash benefit (Sjukpenning, SP), Sickness and activity compensation (Sjuk- och aktivitetsersättning, SA), Parental benefit (Föräldrapenning, FP) and Assistance allowance (Assistansersättning).

Parental benefit (FP) has an upper claims limit which is easily monitored. As most parents use the benefit up to its limit, errors and fraud in FP are therefore mainly a matter of timing of the claims which has little financial impact. FP is thus of less interest for an additional study.

In the study of another benefit, one should primarily use random audits already made by FK. If such data are insufficient, a special audit program has to be carried out, similar to the program in the IFAU project. A program with about 1 000 random audits should be sufficient.

- An empirical study of some benefit can be carried out in order to measure the effect on behaviour of the recommended audit strategies. A similar study has been made for tax auditing in Appelgren (2008).

10 References

Allingham, M.G., Sandmo, A. (1972): "Income Tax Evasion: A Theoretical Analysis", *Journal of Public Economics,* vol.1, pp. 323-338.

Andreoni, J., Erard, B., Feinstein, J.S. (1998): "Tax Compliance." *Journal of Economic Literature*, Vol. 36, June, pp. 818-860.

Appelgren, L. (2008): "The Effect of Audit Strategy Information on Tax Compliance – An Empirical Study." *eJournal of Tax Research*, Vol. 6, No. 1, pp. 67-81.

Appelgren, L. (2011): "Kontrollstrategi för tillfällig föräldrapenning. ISF Arbetsrapport 2011:2.

Becker, G. S. (1968): "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, Vol. 76, March/April, pp. 169-217.

Blumenthal, M., Christian, C., Slemrod, J. (2001): "Do normative appeals affect tax compliance: Evidence from a controlled experiment in Minnesota", *National Tax Journal,* Vol. 54, pp. 125-138.

Engström, P., Hesselius, P., Persson, M. (2006): Överutnyttjande i tillfällig föräldrapenning för vård av barn. IFAU Rapport 2006:9.

Erard, B., Feinstein, J. S. (1994a): "Honesty and Evasion in the Tax Compliance Game." *RAND Journal of Economics,* Vol. 25, No. 1, pp. 1-19.

Försäkringskassan (2007): Utvärdering av TFP-kontroll via arbetsgivare, förskola och skola.

Hasseldine, J., Hite, P., James, S., Tuomi, M. (2007): "Persuasive Communications: Tax Compliance Enforcement Strategies for Sole Proprietors", *Contemporary Accounting Research,* Vol. 24, No. 1, pp. 171-94.

Molander, P. (2011): Optimering av kontrollinsatser – exemplet tillfällig föräldrapenning. ISF Rapport 2011:9.

Reinganum, J. F., Wilde, L. L. (1986): "Equilibrium Verification and Reporting Policies in a Model of Tax Compliance." *International Economic Review,* Vol. 27, No. 3, pp. 739-760.

SOU 2011:3: Sanktionsavgifter på trygghetsområdet. Betänkande avgivet av Utredningen om ett administrativt sanktionssystem inom trygghetssystemen.

Appendix

The content of the IFAU data files

The Audit file KONTROLLERNA.XLS

An Excel file with about 2400 rows with one audit per row, with the following columns:

County

Id number

Group (A, B or C)

Administrator code

Payment date

Extent % (partial or full absence from work)

Start date

End date

Number of days/hours

Paid amount

Audit result (Correct/incorrect)

Error code

Other cause (text at error code 7 Other)

Answer obtained from school/pre-school (Yes/no)

Number of net days (with regard to extent)

Error amount

Etc

The Payment file VAB.CSV

Text file with approximately 2.1 million rows, one per TFP payment. The file contains claims for the period October 2005 – June 2006 and covers the reference period October 1, 2005 – February 28, 2006, and the audit period March 29 – May 31, 2006. The file is estimated to contain data for about 8.5 months. Note that the time period for the file refers to the time when the claim is made and not to the time of sickness.

Group (A, B, C or D)

Start date

End date

Amount

Extent %

Gross days

Net days

Payment month

Id number adult

Id number child

The Population file POPULATION.TXT

Text file containing about 1.3 million persons who have made TFP claims during the total period. The file contains the following data:

Id number of adult

Group (A, B, C or D)

Sex

Birth year/month

Parish code (includes city and county code)

Country of citizenship code

Sector of employment (text)

Education (text)

Income SEK (denoted p21-p24, not explained)

The Relation file RELATION.CSV

Text file containing about 1.1 million children born 1994-2005.

Group (A, B, C or D)

Birth year/month

Sex

Parish code

Id number child

Id number for up to six related persons (children or carers)

SWEDISH SOCIAL INSURANCE INSPECTORATE

address Box 202, 101 24 Stockholm *street address* Fleminggatan 7 Stockholm Sweden

*phone + 46 (*0)8 58 00 15 00 *fax* 08 58 00 15 90 *e-mail* registrator@inspsf.se *web* www.inspsf.se