Insurance Fraud Estimation: More Evidence from the Quebec Automobile Insurance Industry

by Louis Caron and Georges Dionne

Abstract

This article follows a previous study on insurance fraud in the Quebec automobile insurance industry (Dionne and Belhadji, 1996). Results from that research showed that 3 to 6.4% of all claim payments (excluding those for "glass damage only") contained fraud, representing 28 to 61 million dollars in 1994-1995. This evaluation was a minimum since it was limited to observed fraud only. In this paper, we apply a statistical method to estimate the total fraud level in the industry for the same period. Our results show a multiplicative factor of 3.4% of fraudulent files found in Dionne and Belhadji, which means that total fraud payments ranged from 96.2 to 208.4 million dollars in 1994-1995. Our Best Guess Estimator yields roughly a 10% fraud rate or about 113.5 million dollars. An interesting corollary in this finding is that the claim adjusters who participated to the survey (representing 70% of the market), observed only 1/3 of the potential frauds in the studied closed files. One can interpret this number as an index of efficiency for the entire verification process in the industry. A natural question is: Why is this index of efficiency so low? (JEL : G14, G22, D82)

Keywords: Insurance fraud, Quebec automobile insurance industry, observed fraud, estimated fraud, hidden phenomenon, claim adjusters, count data estimators, robustness.

Résumé

Cet article présente une extension d'une étude sur l'évaluation de la fraude à l'assurance dans l'industrie de l'assurance automobile du Québec (Dionne et Belhadji, 1996). Les résultats de cette recherche indiquaient que 3 à 6,4 % des montants des réclamations payées (excluant ceux pour bris de vitre seulement) étaient frauduleux, soit un total des réclamations variant entre 28 et 61 millions de dollars en 1994-1995. Cette évaluation était un plancher, car elle était limitée à la fraude observée seulement. Dans cette recherche, nous appliquons une méthode statistique qui permet d'évaluer la fraude totale dans l'industrie pour la même période. Nos résultats multiplient par 3,4 le pourcentage de dossiers frauduleux obtenus dans Dionne et Belhadji, ce qui représente un total de réclamations variant de 96,2 à 208,4 millions de dollars en 1994-1995. Notre Meilleur Estimateur génère un taux de fraude de 10 % ou 113,5 millions de dollars. Un corollaire intéressant de ce résultat est que les enquêteurs des entreprises qui ont participé à l'enquête (représentant 70 % du marché) ont observé seulement 1/3 de la fraude potentielle dans les dossiers fermés étudiés. Nous pouvons interpréter ce ratio comme un indice d'efficacité du processus de vérification de l'industrie. Une question naturelle est: Pourquoi cet indice est-il si faible ? (JEL : G14, G22, D82)

Mots clés: Fraude à l'assurance, industrie québécoise de l'assurance automobile, fraude observée, fraude estimée, phénomène dissimulé, enquêteurs, estimateurs de données de comptage, robustesse.

Introduction

This article follows a previous study on insurance fraud by Dionne and Belhadji (1996). It uses the same data bank. Eighteen companies have contributed to the survey of this study, representing 70% of the Quebec automobile insurance market in 1994. Claim adjusters randomly reopened 2,509 closed files, or 2,772 coverages, to evaluate the significance of insurance fraud.

Results from this study showed that 3 to 6.4% of all claim payments contained fraud, representing 28 to 61 million dollars in 1994-1995. This evaluation was a minimum since it was limited to observed fraud only. Their definition of fraud included build-up, opportunistic fraud and planned fraud (see Weisberg and Derrig (1993) for a detailed discussion of different fraud definitions; for recent studies on insurance fraud, see the "References" section).

The objective of this paper is to apply a statistical method to estimate the total fraud level in the industry. From the data, investigators found 19 established fraud cases out of the 2,772 coverages, and 123 suspected cases with a degree ranging on a scale from 1 to 10, where 10 means that the case was suspected of having a probability of being fraudulent close to one.

If one considers that only the coverages with established fraud are actually fraudulent, then one obtains a 0.69% fraud level. One can also think that the established and the suspected cases are all fraudulent with a probability equal to one. This observation yields a 5.1% fraud level for the 2,772 coverages or 5.4% for the 2,454 closed files with complete information. In both cases the assumptions are extreme and are limited to observed fraud.

Another possibility is to start with the assumption that, when fraud is established, these coverages are fraudulent with a probability equal to one. This yields a 0.69% lower bound for fraud. We can also say that suspected coverages are "more likely" to represent fraud than the unsuspected ones. In other words, we can assume that at least "some" of these other coverages contain some fraud.

But one may ask: To what extent does the observed fraud underestimate the real fraud? Are we seeing the whole picture or just the tip of an iceberg? This paper proposes an answer.

In the following sections the methodology used for the estimation process is presented and major problems encountered are discussed. A succeeding section presents some estimation results obtained from the data in Dionne and Belhadji (1996), and the last section will wrap-up the results and interpret them in terms of claim payments for the industry. The main results are interpreted in the concluding section.

Problems and Method

In standard statistical evaluations of a ratio, both the numerator and the denominator are perfectly observable. With some subset of the population one can get a robust estimator of the seeked proportion.

The major problem when we have to evaluate the significance of fraud in a given market, is one of estimation. We cannot find easily a proportion of fraud over all coverages because the numerator of this proportion is hidden information. In other words, we do not know with certainty the value of this numerator even in the sample. Consequently, we have to resort to a <u>count data estimator</u> of some <u>hidden phenomenon</u>. The major statistical problem associated with these estimators is their lack of <u>robustness</u>.

Figure I will help to illustrate the problem. Set F represents total fraud in the market while sets E and S show respectively the established and suspected fraud. Clearly, for "fraud proportion" the cardinal of set F is what we are looking for to be our numerator over total claims.

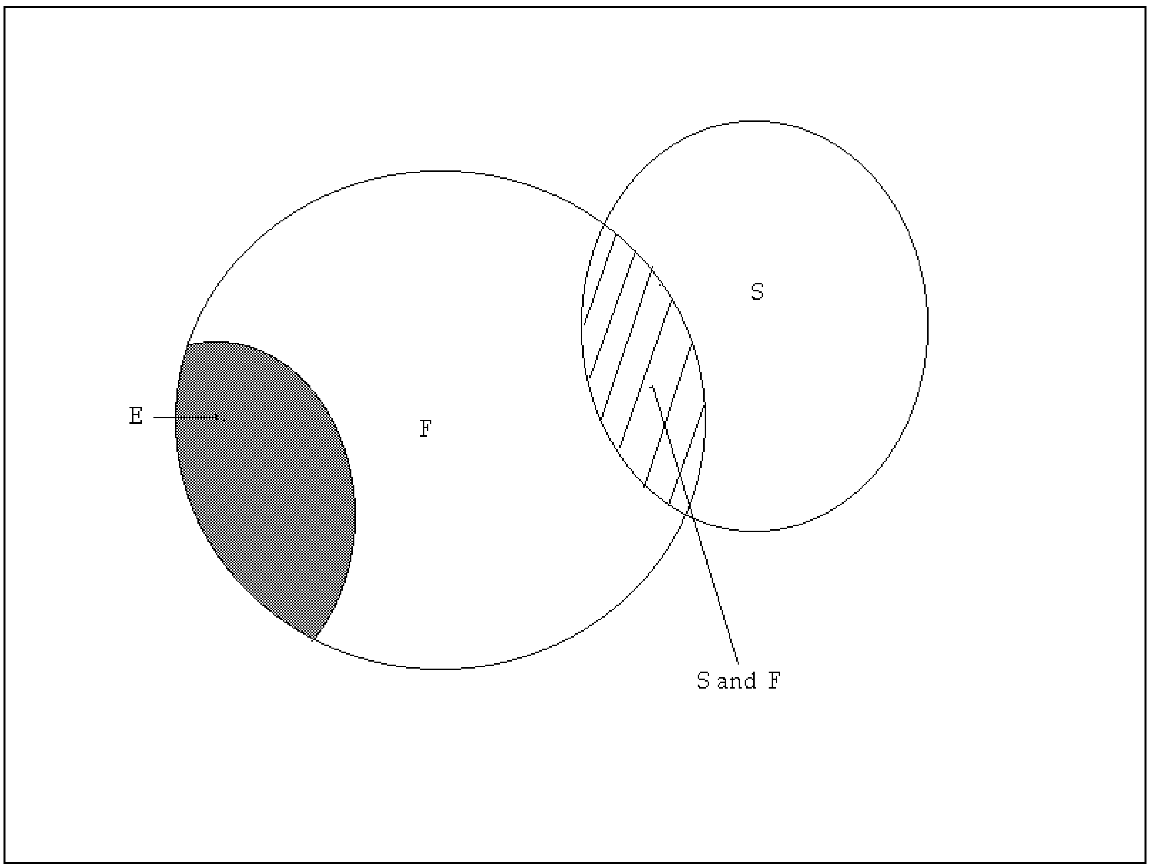

Figure I

The result given in Dionne and Belhadji for observed fraud (19 fraud cases or 0.69%) is represented here by the shaded set E. Since set E is the Established fraud set, then clearly, it is completely contained in the total fraud set F. Set E is in fact a lower bound for set F. Established fraud (set E) is known but it is only part of the total fraud (set F).

We also have a Suspected fraud set (S) with a degree of suspicion for each claim in that set. Some of these suspected fraud cases are really fraudulent, hence they are part of set F, noted here as "S and F". In Dionne and Belhadji (1996), we could see several assumptions on the size of "S and F" ranging from 0 and yielding 0.69% fraud, to 100% of set S and yielding 5.1% fraud. Whatever the assumption used, the total fraud set F was only composed of Established fraud E plus some part of the Suspected fraud S, with the remainder of set F being empty. In other words, Dionne and Belhadji (1996) assumed that claims that are neither established nor at least suspected fraudulent by claim adjusters are never fraudulent.

In order to estimate the cardinal of set F we have to use a count data estimator. The origins of count data estimators date back to Student with Poisson's law, which is well suited for rare occurrences. The advance in genetic science led Fisher to consider the problem with the Negative Binomial law. The Binomial law, for the purpose at hand, was first considered by Binet (1954) and had two major properties.

The first property is that this law has an implicit lower bound, which is the observed number of success in the data. For example, one cannot obtain "100" from a Binomial law with parameters 70 and p, Bin(70,p), no matter what "p" is. So, if the number of established fraud cases in some set is 15, we cannot assume with the Binomial law to have, say, 10 fraudulent cases in this set.

The second property is one of intuition regarding the definition of "p". If we assume that the number of detected fraud we find in any set follows a Bin(n, p), with "n" being the total number of fraud cases, then, by definition, "p" will be the conditional probability of detecting a fraud, given that the claim is fraudulent. So, if we can estimate this "index of efficiency" of claim adjustment staff, we can also find, as a by-product, the total number of fraud, which is what we are mainly looking for.

For example, if we find that a claim adjuster will detect a fraud, given the claim is fraudulent with a probability of 0.5, then according to the Binomial law, since $E[X] = np$, we should double our findings in order to get "n", the total number of fraudulent cases: $n = E(X)/0.5$ where $E(X)$ is the expected number of fraud cases.

Model

Therefore, our assumption is that the detection process of fraud follows a Bin(n,p), with "n" and "p" being the unknown parameters. A method to estimate these parameters is the Method of Moments.

Since there are two parameters in this estimation process, one then needs at least two moments, E[X] and Var[X]. Since our objective is to compute a variance between each group; consequently we need more than one group. For that reason, one has to use a stochastic process to put the data into a number of sets S_1, S_2, ... , S_K.

There is a trade-off in the choice of the number of sets (K). When K is large, the moments are more stable and precise. But as K increases, it becomes more difficult to maintain the Binomial assumption that each set has the same Bin(n,p). The "p" parameter does not change, but the more groups we have, the less elements we have in each group and hence, the less we can say that there is the same number "n" of total fraud cases in each group.

We therefore have to choose a K, or repeat the same experiment with different values of K, and verify how large the variations are between the results for different K's. We will further comment on this point in the next section.

Once we have chosen the number of groups, we can proceed with the estimation of the two moments, and then find the estimation of the two parameters, "n" and "p", as follows.

Let us use the notation μ and σ^2 for the mean and the variance, respectively:

μ = E[X] = np,

σ^2 = V[X] = np(1–p).

Then, we can easily find that $n = \mu^2/(\mu-\sigma^2)$.

However, a major problem arises. As we can see, when $\mu \to \sigma^2$ then $n \to \infty$. This estimator is <u>not robust</u>, which means that little variations in the data will lead to big changes in the estimation. For that reason, we have to use a process to stabilize the estimation. The process used was found and described by Olkin, Petkau and Zidek (1981). Their estimator from the method of moments solves:

$\mathrm{Max}\{\sigma^2\psi^2/(\psi-1), X_{max}\}$

where,

$\psi = \mu/\sigma^2$, when $\mu/\sigma^2 \geq 1 + 1/\sqrt{2}$

$= \max\{z/\sigma^2, 1+\sqrt{2}\}$, otherwise

and,

$z = (X_{max}-\mu)/\sigma$.

Results

We first present the results of one experiment done with six sets of 462 coverages. This experiment represents the average of a thousand estimations with the method described above. Each estimation is not stable, but when we take an average of a hundred or so, the results become much more reliable.

	Occurrence		**Estimation**		
	(n)	**(n/N)%**	**(n)**	**(n/N)%**	**(p)**
E	19	0.6852	10.9449	2.3690	0.2893
E+(S>9)	38	1.3704	22.1362	4.7914	0.2861
E+(S>8)	48	1.7310	27.7267	6.0015	0.2885
E+(S>7)	62	2.2358	35.449	7.6729	0.2915
E+(S>6)	71	2.5604	41.9	9.0693	0.2824
E+(S>5)	78	2.8128	43.7889	9.4781	0.2969
E+(S>4)	100	3.6062	56.4136	12.2107	0.2954
E+(S>3)	108	3.8947	59.4966	12.8781	0.3025
E+(S>0)	127	4.5799	69.5356	15.0510	0.3044
E+S	142	5.1208	76.3894	16.5345	0.3098

Table I: Results with Six Sets (N = 462)

The first column represents the *detection assumption*, which is: "What do we consider as detected fraud?". In terms of figure I, "**E**" represents the entire set E or the Established fraud set. "**S**" stands for the Suspected fraud set or that portion of the set which is calculated as fraud detection. Again, in terms of Figure I, **S** gives the part of set S that is included in set F, or the proportion of set (S and F) over set S. Therefore, the detected fraud set will be set E plus set (S and F).

For example, "**E+(S>4)**" means that set (S and F) is composed of all the suspected fraud cases that have a suspicion degree higher than 4 in the data set of Dionne and Belhadji (1996). That degree of suspicion was included in the data bank and was given by claim adjusters as a "probability of being fraudulent". Hence, in this example, "**E+(S>4)**" means that for this *detection assumption,* we calculate as detected fraud cases as follows: Set E, entirely, plus all suspected cases with a "probability of being fraudulent" equal to 0.5 and higher.

The *detection assumption* ranges from "**E**", the more Optimistic one, found in the first row, where only Established fraud cases are considered as Detected fraud cases, to **"E+S"**, the more pessimistic one, under which all cases of either Established or Suspected fraud are considered as Detected fraud cases.

The second and third columns present the results of the claim adjusters as taken directly from the data bank, and the percentages of fraud proportion. These figures are the observed cardinals of set E plus set (S and F). For example, in the first *detection assumption*, only 19 cases in the data bank were Established as fraudulent, which yields a 0.69% fraud proportion (19/2,772). These two columns show the results presented in Dionne and Belhadji (1996).

The last three columns present the results from the estimation part of this study. The fourth one gives the average "n" estimated over a thousand iterations of the process described earlier. In the fifth column, we can read the fraud proportion obtained where N = 462. Finally, the last one presents the estimated conditional probability of detecting fraud, given there is fraud. In other words, if we give a fraudulent claim to a claim adjuster, then "p" represents the probability that he will detect it as being fraudulent. This, of course, is dependent of the detection assumption.

One important thing to note here is that, as we become more pessimistic in our detection assumption, "p" increases. This is coherent with the intuition that the more we include Suspected fraud cases as Detected, the more we effectively detect fraud. So, as we increase the number of detected fraud with the suspected frauds, the estimated fraud proportion increases, but not linearly so. The relation is increasing but concave.

The percentage of fraud estimated ranges from 2.37% to 16.53% depending on the "optimism degree" in the assumption. This is quite a large bracket but both assumptions are quite extreme. The first one assumes that Suspected cases have no more chances of being fraudulent, and the latter assumes that all Supected cases can be seen as cases where fraud is detected.

If we want a "realistic" fraud estimation, we may consider the Detection assumption to be halfway between the two extremes, which is "E+(S>5)". Here we consider as Detected fraud cases , the Established cases together with Suspected cases where the degree of suspicion, as recorded by the claim adjusters, exceeds five. This gives an estimated "n" of 43.79 fraud cases per set, which yields to an estimation of roughly 9.5% fraudulent claims. We name this assumption the "Best Guess Assumption".

The conditional probability "p" of detecting fraud given there is fraud, under the Best Guess Assumption, was estimated as 0.3. We can compound a *multiplicative factor* with these results. This *multiplicative factor* is defined by the size of the estimation when compared to the observation. In terms of Figure I the *multiplicative factor* is the number of times set E plus set (S and F) enters in the total fraud set (F).

The corresponding *multiplicative factor*, for the Best Guess Assumption, is 3.4. This means that the observed fraud rates, the fraud rates given in the study of Dionne and Belhadji (1996), are multiplied by 3.4 in this study in order to get the real estimated fraud rate in that market.

As we have said earlier, the number of sets (K) was chosen somewhat arbitrarily, with the exception that one had to consider the trade-off in so choosing it. Hence, we only know that the number of sets "K" cannot be close to the value of one or "too large". So we have repeated the same experiment, with a thousand iterations, for different values of "K" (from 5 to 18). In table II we can see some of the results for the estimated percentages.

	Number of Sets (K)					
Assump.	**5**	**6**	**7**	**9**	**11**	**18**
E	2.2860	2.3690	2.3858	2.5819	2.5501	2.7077
E+(S>9)	4.5631	4.7914	4.9345	5.1174	4.9756	5.2368
E+(S>8)	5.7705	6.0015	6.0668	6.4005	6.2785	6.6488
E+(S>7)	7.2835	7.6729	7.5528	8.0999	8.0044	8.5361
E+(S>6)	8.3811	9.0693	9.2734	9.1558	9.1621	9.7969
E+(S>5)	9.2065	9.4781	9.8199	9.9097	10.060	10.762
E+(S>4)	11.724	12.211	12.128	12.494	12.847	13.639
E+(S>3)	12.314	12.878	13.148	13.413	13.966	14.709
E+(S>0)	14.883	15.051	15.608	15.692	16.349	17.222
E+S	16.563	16.535	17.721	17.618	18.014	19.012

Table II: Results for (n/N) by Number of Sets

As we can see, for different numbers of sets, the estimated percentage of fraudulent claims is quite stable. That is, the estimated number of fraudulent claims "n" decreases significantly when we increase the number of sets, but this effect is offset by the decreasing number of claims in each set.

For the Best Guess Assumption "E+(S>5)", the variation in the estimated percentage ranges from 9.2% to 10.8%, which gives us a 10±0.8% interval where we can find the estimated fraud percentage under this assumption.

New monetary estimates for Quebec automobile insurance industry

In this section, we first use the Pessimistic Assumption that sets the degree of suspected fraud at 100%, which means that the total claim payments by the industry for these suspected cases represent detected fraud. We also assume that the multiplicative factor (3.4), obtained from our Best Guess Assumption for the 2,772 coverages, applies to the 2,454 claims for which information on claim payments is available.

Under these assumptions, the total number of fraudulent claims represent 18.4% (5.4% $\times$ 3.4) of total claims and 21.8% (6.4% $\times$ 3.4) of total claim payments, (which amounts to 957,902,484 million dollars in 1994-1995 when excluding "glass damages only"). This yields 208.4 million dollars compared to the 61.3 million obtained in Dionne and Belhadji (1996).

If we now apply the residual monetary amounts for fraud payments obtained from the questionnaire (Realist Assumption #1 in Dionne and Belhadji's study), the residual fraud is equal to 96.2 million instead of 28.3 million or 10% instead of 3% of total claim payments.

Finally, if we restrict the percentage of fraud cases to that of our Best Guess Assumption (E + (S+5)) which means that the fraud rate is 10%, but apply the monetary amounts of the Pessimistic Assumption, we obtain that fraud payments represent 11.85% of total claim payments or 113.5 million dollars of 957,902,484 million dollars.

Conclusion

Our Best Guess Estimator roughly yields a 10% fraud rate, and this result is found to be quite stable. However, the fraud rate is found to have a 16.5% upper bound. The findings in Dionne-Belhadji (1996) are multiplied by 3.4, which were given as a floor estimate, or observed fraud rates. In monetary values, this means that total fraud payments by the industry in 1994-1995, ranged from 96.2 to 208.4 million dollars instead of 28.4 to 61.3 million dollars. In other words, our results indicate that 10 to 21.8% of all claim payments are fraudulent instead of 3 to 6.4%.

An interesting corollary of the present study is the finding that "p" is equal to roughly 1/3. Again "p" is the conditional probability for claim adjustment staff to detect fraud, given the claim is fraudulent. This can be seen as a significantly low index of efficiency for the entire verification process. An important question therefore arises: Why is this index of efficiency so low?

There are countless answers to that question: 1. It can reflect the incompetence of claim adjustment staff to efficiently identify fraud cases. They may not have the adequate experience or training to detect fraud, which in fact is not necessarily their main preoccupation. 2. It can yield serious doubts about the relevance of fraud indicators used to flag possible fraud cases. 3. It may be related to the low quality or quantity of investigations. 4. The results can also reflect an induced laxity by insurers because of the low anticipated benefits of fighting fraud. Choosing the right answer cannot be made without a proper study of the real incentives of each participant in the market to fight against fraud.

In Dionne and Belhadji's study, they found that a large proportion of the fraud cases (93%) were not prosecuted. The main reason for nonprosecution was found to be due to "insufficient proof" (59%). This high percentage of unprosecuted claims for that particular reason naturally triggers a question. Why was the investigation not pushed further?

A possible answer may reside in the fact that many of these claims represent low monetary values. If the claim amount is too small to justify the costs of further investigations, then maybe higher deductibles are in order. Higher deductibles would raise claim levels to the point where investigations could be worth pursuing for the insurers. However, higher deductibles may also increase the benefits of build-up by insureds.

Investigations and prosecutions have also been seen as bad publicity for the investigating and prosecuting firms. The fraud problem is not only a problem of robbery but endangers the very principle of insurance. The question remains with the industry. As researchers, we will focus our attention on finding some statistical and management tools in order to isolate the main causes and improve the claims-premiums ratio in that market.

References

Automobile Insurance Fraud Study (1991), Florida Insurance Research Center, University of Florida, USA.

Binet, F.E. (1954), "The Fitting of the Positive Binomial Distribution When Both Parameters are Estimated From the Sample", *Annals of Eugenics*, vol. 18, p. 117-119.

Caron, L. (1996), "La fraude à l'assurance", Publication Centre de Recherche sur les Transports, Université de Montréal, 60 pages.

Crocker, K.J. and S. Tennyson (1996), "Contracting with Costly State Falsification: Theory and Empirical Results from Automobile Insurance", Working Paper, University of Pennsylvania.

Crocker, K.J. and J. Morgan (1995), "The Optimality of Costly State Falsification: Sharecropping and Insurance Fraud", Working Paper #4-95-9, Department of Economics, The Pennsylvania State University, April.

Cummins, J.D. and S. Tennyson (1992), "Controlling Automobile Insurance Costs", *Journal of Economic Perspectives* 6, Spring, p. 95-115.

Cummins, J.D. and S. Tennyson (1994), "The Tort System 'Lottery' and Insurance Fraud: Theory and Evidence from Automobile Insurance", University of Pennsylvania Working Paper.

Derrig, R.A., H.I. Weisberg and X. Chen (1994), "Behavioral Factor and Lotteries Under No-Fault with a Monetary Threshold: A Study of Massachusetts Automobile Claims", *Journal of Risk and Insurance* 61, June, p. 245-275.

Dionne, G. and B. Belhadji (1996), "Évaluation de la fraude à l'assurance automobile au Québec", *Assurances*, October, vol. 64, p. 365-394.

Dionne, G. and P. St-Michel (1991), "Workers' Compensation and Moral Hazard", *Review of Economics and Statistics* 73, May, p. 236-244.

Dionne, G., P. St-Michel and C. Vanasse (1995), "Moral Hazard, Optimal Auditing and Workers Compensation", *Research in Canadian Workers' Compensation*, T. Thomason and R.P. Chaykowski (Editors).

Olkin, I., A.J. Petkau and J.V. Zidek (1981), "A Comparison of n Estimators for the Binomial Distribution", *Journal of the American Statistical Association,* vol. 76, p. 637-642.

Picard, P. (1994), "Auditing Claims in Insurance Markets with Fraud: The Credibility Issue", CEPREMAP Working Paper.

Weisberg, H.I. and R.A. Derrig (1993), "Quantitative Methods for Detecting Fraudulent Automobile Bodily Injury Claims", Automobile Insurers Bureau of Massachusetts, Boston, 32 pages.