Quantification of losses caused by delinquency in Ecuador

Yannira Chávez[1,2] , Patricia Cortez[3,4] , Paúl Medina [5,6]
[1,3]National Institute of Statistics and Censuses, Quito, Ecuador
[5]Army Polytechnic School, Quito, Ecuador

[2]Corresponding author: Yannira Chávez, e-mail: yannira_chavez@inec.gob.ec
[4]Corresponding author: Patricia Cortez, e-mail: patricia_cortez@inec.gob.ec
[6]Corresponding author: Paúl Medina, e-mail: plmedina@espe.edu.ec

Abstract

This paper quantifies the economic cost of several crimes in Ecuador. In particular, we looked at thefts form vehicle, car accessories and homes, robbery of individual, fraud and intimidation (threat and injuries). To do so, we developed a model to estimate expected and unexpected losses, using Loss Distribution Approach (LDA) methodology. We use data from the Survey of Victimization and Perceptions of Insecurity 2011, carried out by the *National Institute of Statistics and Censuses* (INEC, Spanish acronym) and Administrative Records from the *National Police General Direction of Operations*.

Keywords: cost, economic impact, Loss Distribution Approach.

1 Introduction

Criminal acts and their physical, psychological and economic consequence are a widespread problem in the society, particularly in developing countries such as Ecuador.
For instance, Capa and Gallardo [3] show that 12.70 % of the people over the age 16 were victims of at least one of the following crimes: robbery with force, theft without force, attacks and threats. This implies that more than one of every 10 people are or have been victims of a crime. Among households 16.10 % were victims of at least one of these crimes: theft from housing and/or cars in the year 2008. Moreover during 2011 have registered in 17.26 % of the people over the age 16 were victims of assault and 14.7 % of the families have suffered home burglary or auto theft. Based on these brief Statistics, is not a surprise the fact that residents feel insecure as an evil social, it negatively affects his life and, therefore, demand a greater both security personnel and their property.

Due to the lack and quality of information, it is very hard to determine an accurate estimation of the economic cost of crimes, while in 2011, some of the non-reporting indexes are assault without injuries 84.3 %, home burglary 75.6 %, auto theft 15.5 %, car burglary 81.5 %. On the one hand, many of them are never reported. In 2008, 76.60 % robbery with force, 95.60 % theft without force, 72.70 % theft from housing, theft from cars 12.80 % and 85.60 % theft from accessories were never reported [3], while in 2011, some of the non-reporting indexes are assault without injuries 84.3 %, home burglary 75.6 %, auto theft 15.5 %, car burglary 81.5 %. On the other, information on reported crimes is not always accurate; it is very likely that the information given by individual victim of a crime is not true which makes very difficult to quantify the cost associated to a crime.

Taking into account the problems with the data we estimated the total amount of monetary losses (expected and unexpected) caused by crime based on the theory of Operational Risk (OR). It is important to take into consideration that in finance we talk about, when we talk about risk we refer to the possibility of having losses due to changes in the main factors that affect an asset's value which is the same idea behind our analysis.

In particular, we use the Loss Distribution Approach (LDA) to evaluate the OR. The assumptions behind this methodology are *a)* the total loss a random sum of individual losses, and *b)* all losses are the result of two independent sources of randomness: *frequency* and *severity*.

To estimate our model we combine data from a crime survey and administrative records. The former comes from Survey on Victimization and Perception of Insecurity 2011 conducted by Ecuadors National Institute of Statistics and Census (INEC, spanish acronym) which has data on crimes severity. The latter comes from The National Police General Direction of Operations where we find data on crimes frequency data.

The rest of the document is organizar as follows: Section 2 makes analyzes the theoretical framework used on this paper.Section 3 presents discusses the sources of information and the limitations with the data. Section 4 shows the results of our model. Section 5 concludes.

2 Theoretical Framework

In Finance, any analysis operational risk requires three basic steps: identification, quantification and management. Quantification uses mathematical modeling as its primary tool using aggregate losses distribution method (LDA) as its main technique. An explanation of a variety of techniques for measuring the OR can be found in [4].

Table 1 show a practical application of LDA to crime, the entity could be replaced by state/government or society, business line by *crime* and event risk the type of crime.

Definition	**Traditional – OR**	**Application – OR**
Entity	Bank	State / Government/ Society
Business line	Investment	Crime
Risk Event	Exchange rate	Auto theft, Auto parts theft, Housing theft, Theft from persons, fraud intimidation and injury

Table 1: Technique considerations about LDA application.

2.1. The LDA technique

LDA model is a statistical technique inherited from actuarial field, in which distribution function of aggregate losses is estimated by the convolution of the frequency, a discrete stochastic process and, the severity and a continuous stochastic process. In particular, we first estimated the frequency distribution, and then the severity. The next step is to obtain the joint distribution of aggregate losses from both distributions.
To estimate regulatory capital we apply the definition used on Value at Risk (VaR) in the operational risk context which is known in the literature as OpVaR (Operational Value at Risk). The OpVaR represets a percentile of the loss distribution. Figure 1 shows the practical application of this technique, the *expected loss* is the average loss is expected to occur in a given period and the *unexpected loss* consist of the difference between OpVaR and expected loss.

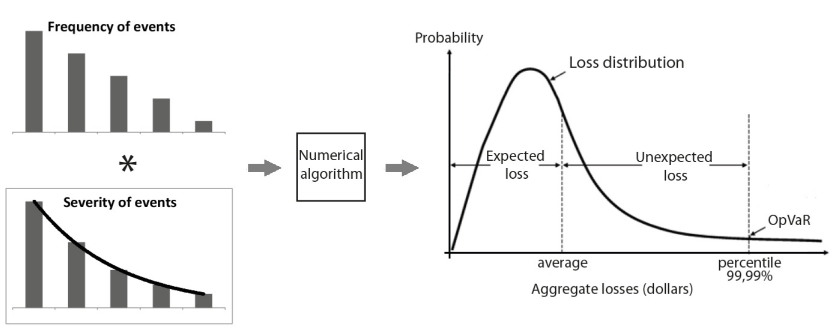

Figure 1: LDA model. Source: Prepared by the authors.

Formally, the aggregate loss for an event is given by:

$$S = X_1 + X_2 + \cdots + X_N,$$

where N, is a counting random variable and X_i is the random variable corresponding to the severity of event occurrence. The LDA model considers the following assumptions for every risk class: *a)* the frequency is a random variable which is independent of the severity, and *b)* the losses observations are homogeneous, independent and identically distributed. Under these assumptions, the aggregate loss distribution function can be presented as

$$G_s(x) = P(S \leq x) = \sum_{n=0}^{\infty} P(N = n) F_X^{*n}(x). \tag{1}$$

where the *n–convolution*, denoted by $F_X^{*n}(x)$, is given by

$$P(X_1 + X_2 + \cdots + X_n \leq x) = F_X * F_X * \cdots * F_X = F_X^{*n}(x). \tag{2}$$

with

$$F_X(x) = P(X \leq x).$$

The assumption of the model (see above) show that *a)* admitted that the frequency and severity are two independent sources of randomness. The assumption *b)* , means that two different losses inside the same class are homogeneous, independent and identically distributed.
It is quite hard to obtain an analytical solution of equation (1), therefore, we used numerical methods such as Monte Carlo simulation and Panjer recursive approach to solve this problem. Other tools like the approximate fast Fourier transform can be used as well (see e.g. [4]). For this paper we choose to used Monte Carlo simulation to estimate the aggregate loss function of crime, because it allow us to estimate the distribution using a large number of randomly generated scenarios from the severity and frequency distributions.

2.2. Frequency and severity distributions

Frequency is the random variable that represents the number of events which produce losses in an interval. It followa a reference probability distribution. For instance [4, 7] show scenarios for the Poisson distribution and its applications to real situations. However, it also recommended to consider alternatives such as Negative Binomial, Binomial, Geometric and Hypergeometric distributions.

Severity is the random variable that represents the impact of the event in terms of economic loss and follows a probability distribution reference. [4, 7] show lognormal distribution and its applications. As before, other distributions should be consider such as Weibull, Pareto and Exponential.

3 Data

We use data from a Survey on Victimization and Perception of Insecurity 2011 conducted by INEC which has data on severity values as the average of the crime-provoked financial losses reported by the respondents. We also uses information from The National Police General Direction of Operations administrative records which contents data on crimes frequency. A brief summary of the data is shown on Table 2. We can see that frequency and severity of crime are concentrated robbery and auto theft, repectively.

We uses these two sources of information because our model required that data on frequency and severity must be two independent sources of randomness. This assumption is met because data coming from INEC and the Police administrative records are collected with different approaches.

Risk	Frequency	Severity
Auto theft	5 908	$ 4825,19
Car burglary	5 782	$ 579,74
Home burglary	13 482	$ 1 380,77
Robbery	21 107	$ 369,29
Fraud	10 565	$ 1 430,75
Assault without injuries	11 331	$ 443,46
(Assault with) Injuries	6 684	$ 128,65
Total	**74 859**	**$ 9 157,85**

Table 2: Data frequency and severity estimates. Source: prepared by the authors.

4 Crime estimates that cause lost

To estimate from a theoretical point of view crime-provoked loss, we need to know the theoretical distribution functions, either discrete or continuous, for the frequency and severity, respectively. Because the amount of available data does not allows us apply any robust-enough parametric or non-parametric test, we follow [4, 7] and estimate the loss function using Poisson and Negative Binomial distribution for frequency and Lognormal and Weibull for severity and simulated four possible scenarios: *a)* Negative Binomial - Lognormal, *b)* Negative Binomial - Weibull, *c)* Poisson - Lognormal, *b)* Poisson - Weibull.

Once selected the pair of distributions, we estimate the aggregate losses distribution, LDA, by considering

$$\text{LDA} = Frequency * Severity.$$

Next, we applied the Monte Carlo simulation method in which the absolute error decreases as $\frac{1}{\sqrt{N}}$, where $N \in \mathbb{N}$ is the number simulations.

The results for each scenario for different percentiles are shown in Table 3.[1] In addition, Table 3 includes, a losses interval for the Negative Binomial - Poisson and Lognormal - Weibull simulations.

Negative Binomial - Lognormal			
Percentile	**OpVaR**	**Expected losses**	**Unexpected losses**
99,99 %	$ 99.968.501,25	$ 14.148.719,25	$ 85.819.782,00
99,90 %	$ 65.476.646,63	$ 14.148.719,25	$ 51.327.927,37
99,00 %	$ 44.355.692,60	$ 14.148.719,25	$ 30.206.973,35
Negative Binomial - Weibull			
Percentile	**OpVaR**	**Expected losses**	**Unexpected losses**
99,99 %	$ 57.847.808,66	$ 13.956.203,02	$ 43.891.605,64
99,90 %	$ 48.987.077,79	$ 13.956.203,02	$ 35.030.874,77
99,00 %	$ 37.538.486,19	$ 13.956.203,02	$ 23.582.283,18
Poisson - Lognormal			
Percentile	**OpVaR**	**Expected losses**	**Unexpected losses**
99,99 %	$ 93.837.004,28	$ 14.150.619,45	$ 79.686.384,83
99,90 %	$ 61.191.483,46	$ 14.150.619,45	$ 47.040.864,01
99,00 %	$ 39.588.339,68	$ 14.150.619,45	$ 25.437.720,24
Poisson - Weibull			
Percentile	**OpVaR**	**Expected losses**	**Unexpected losses**
99,99 %	$ 46.566.398,14	$ 13.957.344,18	$ 32.609.053,96
99,90 %	$ 38.996.948,83	$ 13.957.344,18	$ 25.039.604,65
99,00 %	$ 31.291.892,40	$ 13.957.344,18	$ 17.334.548,22

Table 3: Estimation of expected and unexpected losses, considering distribution combinations for frequency and severity. Source: prepared by the authors.

[1] We consider 200000 for each simulation.

Next, we present the total aggregate losses distribution for each scenario.

1. **Negative Binomial - Lognormal**

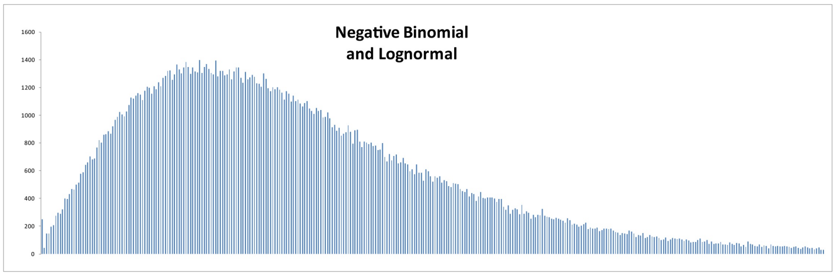

Figure 2: Total aggregate losses distribution. Source: prepared by the authors.

2. **Negative Binomial - Weibull**

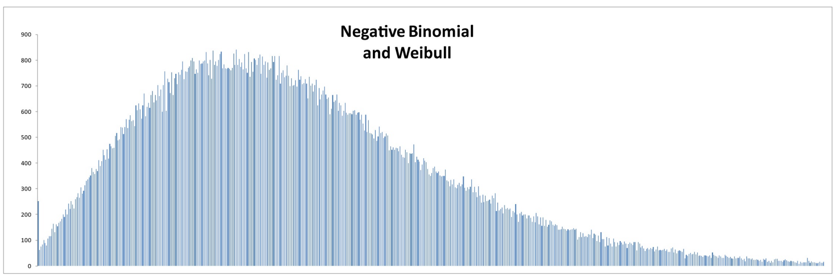

Figure 3: Total aggregate losses distribution. Source: prepared by the authors.

3. **Poisson - Lognormal**

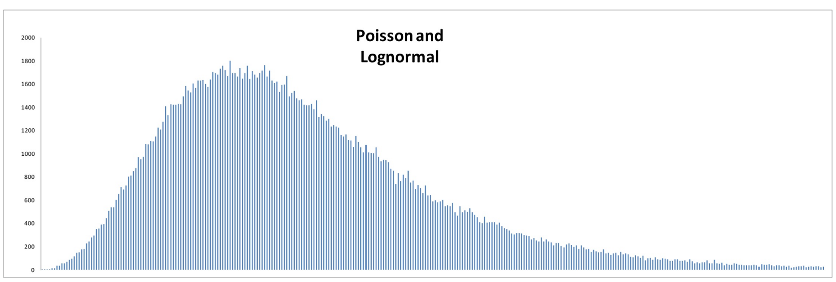

Figure 4: Total aggregate losses distribution. Source: prepared by the authors.

4. **Poisson - Weibull**

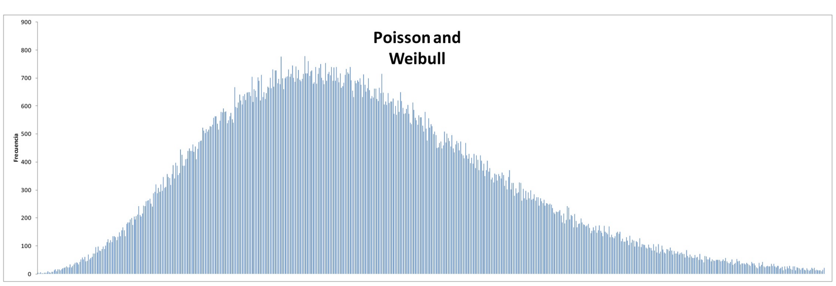

Figure 5: Total aggregate losses distribution. Source: prepared by the authors.

5 Conclusions

In this paper, we estimated the expected and unexpected cost of crime adapting the theory and empirical tools commonly used in operational risk analysis.

We obtained different levels of loss, expected and unexpected, considering various percentiles for each of the four considered scenarios. This is due to limited data, as cannot determine the true probability distribution.

Taking the percentile corresponding to 99.99 %, we can set a crime-provoked losses interval, between \$ 33 and \$ 86 million, approximately. These values would come from the Poisson - Weibull and Negative Binomial - Lognormal scenarios.

Finally, we established that the superior and inferior limits of the crime-provoked losses interval are quite sensitive to the percentile choice. So, if we take e.g. the percentile of 99.00 %, the range of losses would be established between \$ 17 and \$ 33 million, approximately.

References

[1] Aue, F. y Kalkbrener, M. (2006). *LDA at work: Deutsche Bank's approach to quantifying operational risk*, Journal of Operational Risk, 1, 45-93.

[2] Cajas J. (2011). *Modelos de enfoque de Medición Avanzado de Riesgo Operativo (EMA)*, Contribuciones a la Economía, Grupo Eumed.net, Universidad de Ḿlaga.

[3] Capa, H. y Gallardo C. (2008). *Encuesta de Victimización y Percepción*, 1-109.

[4] Cruz, M. (2002). *Modeling, measuring and hedging operational risk*, New York: John Wiley & Sons.

[5] Franco, L. y Murillo, J. (2008). *Loss Distribution Approach (LDA): Metodologá actuarial aplicada al riesgo operacional*, Revista Ingenierías Universidad de Medellín, 7, 143-156.

[6] Nieto, M. (2004). *The modelling of Operational Risk: Experience with the Analysis of the Data Collected by the Basel Committee*, Documento de trabajo, Banco de Italia.

[7] Shevchenko, P. y Wthrich, M. (2006). *The structural modeling of operational risk via Bayesian inference: combining loss data with expert opinions*, Journal of Operational Risk, 1, 3-26.

[8] Claramunt, M. Boj, E. Costa, T. y Mrmol, T. (2012). *Sistemas Informáticos aplicados a la matemática actuarial no vida una propuesta con R*, Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA, 13, 1-26.