TYPES OF CLAIMS IN MOTOR THIRD PARTY LIABILITY INSURANCE

Michel DENUIT¹ and Julien TRUFIN²

¹Institut de Statistique, Biostatistique et Sciences Actuarielles UCL, Louvain-la-Neuve, Belgique

²Département de Mathématique ULB, Bruxelles, Belgique

EAJ Conference, Lyon, September 8, 2016

Introduction

Approach

Modelling claims with rapid settlement

Modelling claims with longer development

Reserve calculations

References

Introduction

Approach

Modelling claims with rapid settlement

Modelling claims with longer development

Reserve calculations

References

Classical approach based on aggregated run-off triangles

Traditional reserving models based only on a few cells in an upper triangle.

	D	evelop	ome	nt
Accident Year		j		
	*	*	*	*
i	*	X_{ij}	*	
	*	*		
	*			

- ► Such a **crude aggregation** of the available data inevitably leads to a huge loss of information.
- This classical approach appears to be somewhat outdated at the "Big Data" era.

Data structure

Event	No	Year	Amount
Occurrence	16,384	2009	-
	20,784	2010	-
Declaration	16,384	2009	-
	20,784	2010	-
Payments	16,384	2009	5,022
	16,384	2010	67,363
	16,384	2011	903
	16,384	2012	6,295
	16,384	2013	13,850
	16,384	2014	0
	20,784	2010	1,605
	20,784	2011	0
Closure	16,384	Not settled	-
	20,784	2011	-

Introduction

Approach

Modelling claims with rapid settlement

Modelling claims with longer development

Reserve calculations

References

Data set:

- Extracted from a motor third party liability insurance portfolio.
- ▶ The **observation period:** calendar years 2004 till 2014.
 - ☐ The available information concerns accident years 2004 to 2014 so that we have observed developments j up to n = 11.
- ► There are 52, 155 claims in the data set. Among them, 4,023 claims are still open at the end of the observation period.

Observed numbers of reported claims:

	1	2	3	4	5	6	7	8	9	10	11
2004	4,022	165	8	1	0	0	0	0	0	0	0
2005	4,190	174	5	1	0	0	0	0	0	0	
2006	4,331	210	2	2	0	0	0	0	0		
2007	4,743	255	9	3	0	0	0	0			
2008	5,046	222	8	1	0	0	0				
2009	5,168	191	10	0	0	0					
2010	4,612	217	7	1	0						
2011	4,394	200	9	1							
2012	4,299	162	7								
2013	4,557	169									
2014	4,753										

Observed numbers of closed claims:

	1	2	3	4	5	6	7	8	9	10	11
2004	2,266	1,509	203	94	42	38	9	14	5	3	4
2005	2,428	1,582	172	62	46	29	17	12	8	4	
2006	2,433	1,607	228	139	65	29	22	5	7		
2007	2,451	1,853	433	136	59	38	14	8			
2008	2,643	2,079	287	141	58	30	21				
2009	2,607	2,105	385	127	71	29					
2010	1,782	2,442	340	146	66						
2011	1,793	2,223	337	123							
2012	1,852	2,017	373								
2013	1,859	2,197									
2014	1,925										

Observed numbers of closed claims:

	Reported	Closed	Difference
2004	4,196	4,187	9
2005	4,370	4,360	10
2006	4,545	4,535	10
2007	5,010	4,992	18

- ▶ n= number of observation periods, $i \in \{1, 2, ..., n\}$ (or $i \in \{2004, 2005, ..., 2004 + n 1 = 2014\}$);
- ω = longest time to development, $j \in \{1, 2, \dots, \omega\}$;
- ▶ Here: $n = 11 < \omega$.

Descriptive statistics for yearly payments:

			2	3	4	5	6	,	8	
2004	Num. pay.	2,848	1,459	236	124	68	39	18	18	
	Mean	1,133	1,877	2,713	4,349	4,446	9,894	16,765	4,422	
2005	Num. pay	3,001	1,492	207	97	53	42	24	21	
	Mean	1,112	1,659	3,168	5,455	5,132	14,882	25,781	8,997	
2006	Num. pay	3,007	1,659	268	117	61	41	21	10	
	Mean	1,164	1,624	5,799	4,494	7,287	6,055	6,141	4,688	
2007	Num. pay	3,246	1,893	322	170	79	48	24	16	
	Mean	1,159	1,905	2,679	3,500	7,401	8,243	12,140	13,148	
2008	Num. pay	3,574	1,816	304	125	71	37	22		
	Mean	1,104	1,720	2,189	4,203	4,611	7,775	6,310		
2009	Num. pay	3,545	1,877	300	131	90	51			
	Mean	1,142	1,919	3,981	4,379	6,896	9,129			
2010	Num. pay	2,874	2,072	338	161	75				
	Mean	1,663	1,984	3,637	5,147	14,935				
2011	Num. pay	2,777	1,930	327	119					
	Mean	1,601	1,982	2,441	5,171					
2012	Num. pay	2,860	1,749	282						
	Mean	1,716	2,328	4,390						
2013	Num. pay	2,924	1,844							
	Mean	1,637	2,230							
2014	Num. pay	2,723								
	Mean	1,662								

Minor losses vs major losses

- ▶ In this MTPL context, we adopt the following definitions:
 - ▶ Minor losses = reported and settled at $j \le \omega_1 = 2$ (\Rightarrow rapidly settled);
 - ▶ Major losses = settled at $j > \omega_1 = 2$.

Introduction

Approach

Modelling claims with rapid settlement

Modelling claims with longer development

Reserve calculations

References

- Claims settled relatively rapidly: usually cheaper than those requiring longer settlement periods.
 - ⇒ We isolate claims that are reported and settled rapidly.
- ▶ **Total payment** X_{ij} in calendar year i+j-1 for claims originating in accident year i and settled in at most $\omega_1=2$ years **is decomposed into the compound sum**

$$X_{ij} = \sum_{k=1}^{N_{ij}^{(o)}} X_{ijk}, \ j \in \{1, 2\},$$

where

- ▶ $N_{ij}^{(o)}$ = number of claims with short development originating in accident year i, reported at or before development j, still open at development j;
- ▶ X_{ijk} = total payment (possibly 0) made in calendar year i + j 1 for the kth claim with short development originating in accident year i still open at development j.

Number of reported/closed claims $(N_{ij}^{(r)} \text{ and } N_{ij}^{(c)})$

Notations:

- ▶ $N_{ij}^{(r)}$ = number of claims with short development that occurred in accident year i and were reported to the insurer at development j (i.e. during calendar year i + j 1);
- ▶ $N_{ij}^{(c)}$ = number of claims with rapid settlement originating in accident year i that were reported at development j or before and closed during calendar year i + j 1.

Number of reported/closed claims $(N_{ij}^{(r)} \text{ and } N_{ij}^{(c)})$

▶ Specification:

We use the following specifications (in line with Chain-Ladder)

$$\mathsf{E}[\mathit{N}_{ij}^{(r)}] = \alpha_i \beta_j^{(r)} \text{ and } \mathsf{E}[\mathit{N}_{ij}^{(c)}] = \alpha_i \beta_j^{(c)}$$

subject to the usual identifiability constraints

$$\sum_{j=1}^{\omega_1} \beta_j^{(r)} = \sum_{j=1}^{\omega_1} \beta_j^{(c)} = 1.$$

► This ensures that the total number

$$N_i = \sum_{i=1}^{\omega_1} N_{ij}^{(r)} = \sum_{i=1}^{\omega_1} N_{ij}^{(c)}$$

of claims with short development originating in accident year i has mean $E[N_i] = \alpha_i$. Also, we have

$$\beta_j^{(r)}$$
 = probability that a minor claim is reported at lag j ; $\beta_j^{(c)}$ = probability that a minor claim is closed at lag j .

Number of reported/closed claims $(N_{ii}^{(r)} \text{ and } N_{ii}^{(c)})$

- Parameters estimates:
 - ▶ The parameters $(\alpha_i, \beta_i^{(r)}, \beta_i^{(c)})$ are estimated

 - □ From the observed $N_{ij}^{(r)}$ and $N_{ij}^{(c)}$;
 □ By maximum likelihood assuming that the observed counts $N_{ij}^{(r)}$ and $N_{ij}^{(c)}$ are independent and Poisson distributed.

Number of open claims $N_{ij}^{(o)}$

- ► Link between $N_{ij}^{(r)}$, $N_{ij}^{(c)}$ and $N_{ij}^{(o)}$:
 - ► The numbers of open claims can directly be obtained from the numbers of reported and closed claims. Indeed:

$$\sum_{k=1}^{j} N_{ik}^{(r)} = \sum_{k=1}^{j-1} N_{ik}^{(c)} + N_{ij}^{(o)}$$
$$\Rightarrow N_{ij}^{(o)} = \sum_{k=1}^{j} N_{ik}^{(r)} - \sum_{k=1}^{j-1} N_{ik}^{(c)}.$$

▶ Hence, we get

$$\mathsf{E}[N_{ij}^{(o)}] = \alpha_i \left(\sum_{k=1}^j \beta_k^{(r)} - \sum_{k=1}^{j-1} \beta_k^{(c)} \right).$$

Observations and parameter estimates:

i	$N_{i1}^{(r)}$	$N_{i2}^{(r)}$	Total	$N_{i1}^{(c)}$	$N_{i2}^{(c)}$	Total	$\widehat{\alpha}_{i}$
2004	3,641	134	3,775	2,266	1,509	3,775	3,775
2005	3,872	138	4,010	2,428	1,582	4,010	4,010
2006	3,872	168	4,040	2,433	1,607	4,040	4,040
2007	4,127	177	4,304	2,451	1,853	4,304	4,304
2008	4,541	181	4,722	2,643	2,079	4,722	4,722
2009	4,561	151	4,712	2,607	2,105	4,712	4,712
2010	4,065	159	4,224	1,782	2,442	4,224	4,224
2011	3,871	145	4,016	1,793	2,223	4,016	4,016
2012	3,754	115	3,869	1,852	2,017	3,869	3,869
2013	3,934	122	4,056	1,859	2,197	4,056	4,056
$\widehat{\beta}_{j}^{(r)}$	0.964	0.036	$\widehat{\beta}_{j}^{(c)}$	0.451	0.549		

Yearly payments X_{ijk} per open claims

Specification:

- ▶ Yearly payments X_{ijk} can be equal to 0.
 - ⇒ We account for a probability mass at zero, namely

$$P[X_{ijk}=0]=\zeta_j.$$

▶ The conditional mean $E[X_{ijk}|X_{ijk}>0]$ is then given by

$$\mathsf{E}[X_{ijk}|X_{ijk}>0]=\gamma_{i+j-1}\xi_j.$$

- \square γ_{i+j-1} models inflation;
- \square ξ_j models the development effect.

Inflation:

- Is included by means of hedonic regression techniques.
- Can then be modelled and projected to the future, for instance $\widehat{\gamma}_t \sim MA(1)$ model (that can be seen as an AR(1) observed with errors).

Yearly payments X_{ijk} per open claims

Parameters estimates:

- ▶ The parameters $(\zeta_i, \gamma_{i+j-1}, \xi_i)$ are estimated
 - \square From the observed yearly payments X_{ijk} ;
 - By maximum likelihood assuming that the observed payments X_{ijk} are mutually independent and Gamma distributed.

Parameters estimates:

▶ Probability mass at zero:

► Development effect:

► Inflation:

$$\widehat{\gamma}_{2004+I} = (1+\widehat{\gamma})^I$$
 with $\widehat{\gamma} = 1.34\%$.

Introduction

Approach

Modelling claims with rapid settlement

Modelling claims with longer development

Reserve calculations

References

Related development triangles

Notations:

- ▶ $M_{ij}^{(r)}$ = number of claims with long development originating in accident year i, reported at development j, i.e. during calendar year i + j 1;
- ▶ $M_{ij}^{(c)}$ = number of claims with long development originating in accident year i, reported at development j or before and closed during calendar year i + j 1,

$$M_{ij}^{(c)}=0$$
 for $j\leq\omega_1$;

▶ $M_{ij}^{(o)}$ = number of claims with long development originating in accident year i, still open at development j.

Number of reported/closed claims $(M_{ii}^{(r)})$ and $M_{ii}^{(c)}$

- Specification:
 - ▶ We use the following specifications

$$\mathsf{E}[M_{ij}^{(r)}] = \delta_i \theta_j^{(r)} \text{ and } \mathsf{E}[M_{ij}^{(c)}] = \delta_i \theta_j^{(c)}$$

subject to the usual identifiability constraints

$$\sum_{j=1}^{\omega} \theta_j^{(r)} = \sum_{j=1}^{\omega} \theta_j^{(c)} = 1.$$

▶ This ensures that the total number

$$M_i = \sum_{j=1}^{\omega} M_{ij}^{(r)} = \sum_{j=\omega_1+1}^{\omega} M_{ij}^{(c)}$$

of claims with long development originating in accident year i has mean $\mathsf{E}[M_i] = \delta_i$. Also we have

 $\theta_j^{(r)}=$ proba. that a claim with long development is reported at lag j; $\theta_j^{(c)}=$ proba. that a claim with long development is closed at lag j, with $\theta_i^{(c)}=0$ for $j<\omega_1$.

Number of reported/closed claims $(M_{ii}^{(r)})$ and $M_{ii}^{(c)}$

Parameters estimates:

- ▶ The parameters $(\delta_i, \theta_i^{(r)}, \theta_i^{(c)})$ are estimated

 - □ From the observed $M_{ij}^{(r)}$ and $M_{ij}^{(c)}$; □ By maximum likelihood assuming that the observed counts $M_{ii}^{(r)}$ and $M_{ii}^{(c)}$ are **independent** and **Poisson** distributed.

Number of open claims $M_{ij}^{(o)}$

- ▶ Link between $M_{ii}^{(r)}$, $M_{ii}^{(c)}$ and $M_{ii}^{(o)}$:
 - As

$$\sum_{k=1}^{j} M_{ik}^{(r)} = \sum_{k=1}^{j-1} M_{ik}^{(c)} + M_{ij}^{(o)}$$

we then get

$$\mathsf{E}[M_{ij}^{(o)}] = \delta_i \left(\sum_{k=1}^j \theta_k^{(r)} - \sum_{k=1}^{j-1} \theta_k^{(c)} \right).$$

▶ Observed number of reported claims $M_{ij}^{(r)}$ and parameter estimates:

i	1	2	3	4	5-11	Total	$\widehat{\delta}_i$
2004	381	31	8	1	0	421	421
2005	318	36	5	1	0	360	360
2006	459	42	2	2	0	505	505
2007	616	78	9	3	0	706	706
2008	505	41	8	1	0	555	555
2009	607	40	10	0	0	657	657
2010	547	58	7	1	0	613	613
2011	523	55	9	1		588	588
2012	545	47	7			599	600
2013	623	47				670	680
$\widehat{\theta}_{j}^{(r)}$	0.901	0.084	0.013	0.002	0		

▶ Observed number of closed claims $M_{ij}^{(c)}$ and parameter estimates:

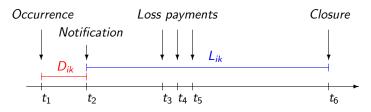
	1	2	3	4	5	6	7	8	9	10	11
2004	0	0	203	94	42	38	9	14	5	3	4
2005	0	0	172	62	46	29	17	12	8	4	
2006	0	0	228	139	65	29	22	5	7		
2007	0	0	433	136	59	38	14	8			
2008	0	0	287	141	58	30	21				
2009	0	0	385	127	71	29					
2010	0	0	340	146	66						
2011	0	0	337	123							
2012	0	0	373								
$\widehat{\theta}_{j}^{(c)}$	0	0	0.522	0.212	0.103	0.059	0.032	0.019	0.015	0.009	0.010

▶ **Remark:** The tail factors $\widehat{\theta}_{12}^{(c)}$ and $\widehat{\theta}_{13}^{(c)}$ have been set to

$$\widehat{\theta}_{11}^{(c)} = \widehat{\theta}_{12}^{(c)} = \widehat{\theta}_{13}^{(c)} = 1\%.$$

- ► Here, we have **fewer losses with longer developments**.
 - ⇒ We build the loss model from **individual claim information**, before aggregating in yearly total.
- We account for
 - ▶ A random reporting lag $D_{ik} \in \{1, 2, ..., \omega\}$;
 - ▶ A random time to settlement $L_{ik} \in \{0, 1, ..., \omega\}$, equal to 0 if settled during the reporting year.

Graphically:



- ▶ The **pair** (D_{ik}, L_{ik}) is modelled as follows.
 - ▶ The marginal distribution of the reporting lag D_{ik} is given by

$$P[D_{ik}=d]=\theta_d^{(r)}.$$

As D_{ik} and L_{ik} are correlated, we specify **the conditional** distribution of L_{ik} given $D_{ik} = d$.

- ▶ The **yearly payments** $Y_{i,k,D_{ik}+h}$, $h = 0,...,L_{ik}$ are modelled as follows:
 - Discrete mixture with three components:
 - \Box a **lighter-tailed component** with probability τ_h such as Gamma or Inverse Gaussian distributions;
 - \square a **heavier-tailed component** with probability ρ_h with Pareto type 2 distribution;
 - □ as well as a probability mass at zero

$$P[Y_{i,k,D_{ik}+h}=0]=1-\tau_h-\rho_h.$$

► The average payment is

$$\chi_{1,h}(1+g_1)^{i+D_{ik}+h-2}$$

for the Gamma component and

$$\chi_{2,h}(1+g_2)^{i+D_{ik}+h-2}$$

for the Pareto type 2 component.

Parameter estimates:

ŀ	ר	0	1	2	3	4	5	6
$\hat{ au}$	h	0.170	0.262	0.153	0.063	0.000	0.000	
$\widehat{ ho}$	h	0.193	0.293	0.355	0.450	0.549	0.564	
$1-\widehat{ au}_h$	$-\widehat{\rho}_h$	0.637	0.445	0.492	0.487	0.451	0.436	
ŀ	ר	7	8	9	10			
$\hat{ au}$	h	0.000	0.000	0.000	0.000			
$\widehat{ ho}$	h	0.545	0.576	0.607	0.500			
$1-\widehat{ au}_h$	$-\widehat{\rho}_h$	0.455	0.424	0.393	0.500			
h	0	1	2		3	4	5	6
Ŷ1 b	3.432	3.765	1.903	3	652	_	_	-

h	0	1	2	3	4	5	6
$\widehat{\chi}_{1,h}$	3,432	3,765	1,903	652	-	-	-
$\widehat{\chi}_{2,h}$	3,061	4,019	4,154	5,318	6,011	8,253	9,709
h	7	8	9	10			
$\widehat{\chi}_{1,h}$	-	-	-	-			
$\widehat{\chi}_{2,h}$	10,167	10,145	10,231	10,178			

The estimated inflation rates g_1 and g_2 are $\widehat{g}_1=0.03\%$ and $\widehat{g}_2=1.96\%$

Introduction

Approach

Modelling claims with rapid settlement

Modelling claims with longer development

Reserve calculations

References

Reserve for minor losses

▶ For accident years $i \in \{n - \omega_1 + 2, ..., n\}$, the expected value of the outstanding claims with short development is

$$E\left[\sum_{j=n-i+2}^{\omega_{1}} X_{ij}\right] = \sum_{j=n-i+2}^{\omega_{1}} E\left[N_{ij}^{(o)}\right] E\left[X_{ij1}\right]$$

$$= \sum_{j=n-i+2}^{\omega_{1}} \alpha_{i} \left(\sum_{k=1}^{j} \beta_{k}^{(r)} - \sum_{k=1}^{j-1} \beta_{k}^{(c)}\right) (1 - \zeta_{j}) \gamma_{i+j-1} \xi_{j}.$$

• $\omega_1 = 2 \Rightarrow$ only the last accident year 2014 is concerned. Hence, for i = 2014, we get

$$E[X_{i2}] = \alpha_i \left(\sum_{k=1}^2 \beta_k^{(r)} - \beta_1^{(c)} \right) (1 - \zeta_2) \gamma_{i+1} \xi_2$$
$$= \alpha_i \left(1 - \beta_1^{(c)} \right) (1 - \zeta_2) (1 + \gamma)^{11} \xi_2.$$

▶ **Remark:** The parameter α_i for i = 2014 still needs to be estimated. Since

$$\mathsf{E}\left[\mathsf{N}_{i1}^{(c)}\right] = \alpha_i \beta_1^{(c)},$$

we can estimate α_i by $N_{i1}^{(c)}/\hat{\beta}_1^{(c)} = 1,925/0.451 = 4,271$.

▶ Hence, the **reserve for claims with short development** is

$$\hat{E}[X_{i2}] = 3,040,309.$$

Reserve for major losses

- ▶ For all accident years (except the first one if $n = \omega$), we must add the reserve for claims with longer development.
- ▶ If only total amounts matter, we can aggregate the payments of all major losses to get

$$Z_{ij} = \sum_{k=1}^{M_{ij}^{(o)}} Z_{ijk}$$

where the Z_{ijk} are independent and identically distributed, and independent of $M_{ij}^{(o)}$.

Reserve for major losses

▶ The **distribution of** Z_{ijk} can be obtained as a mixture by conditioning with respect to D_{ik} . Precisely,

$$P[Z_{ijk} \le z] = \sum_{d=1}^{j} P[Z_{ijk} \le z | D_{ik} = d] P[D_{ik} = d | D_{ik} \le j]$$

where

$$P[Z_{ijk} \le z | D_{ik} = d] = P[Y_{i,k,d+(j-d)} \le z]$$

and

$$P[D_{ik} = d | D_{ik} \le j] = \frac{\theta_d^{(r)}}{\theta_1^{(r)} + \ldots + \theta_j^{(r)}}.$$

Reserve for major losses

► Then, **for accident year** *i*, the expected value of the outstanding claims with long development is

$$\mathsf{E}\left[\sum_{j=n-i+2}^{\omega} Z_{ij}\right] = \sum_{j=n-i+2}^{\omega} \mathsf{E}\left[M_{ij}^{(o)}\right] \mathsf{E}\left[Z_{ij1}\right]$$

with

$$E[Z_{ij1}] = \sum_{d=1}^{j} E[Y_{i,1,d+(j-d)}] \frac{\theta_d^{(r)}}{\theta_1^{(r)} + \ldots + \theta_j^{(r)}}$$

and

$$\mathsf{E}[Y_{i,1,d+(j-d)}] = \tau_{j-d}\chi_{1,j-d}(1+g_1)^{i+j-2} + \rho_{j-d}\chi_{2,j-d}(1+g_2)^{i+j-2}.$$

▶ For **the last accident year 2014**, δ_i still need to be estimated. As we know that

$$\mathsf{E}\left[N_{i1}^{(r)}+M_{i1}^{(r)}\right]=\alpha_i\beta_1^{(r)}+\delta_i\theta_1^{(r)},$$

we estimate δ_{2014} by

$$\widehat{\delta}_{2014} = \frac{N_{2014,1}^{(r)} + M_{2014,1}^{(r)} - \widehat{\alpha}_{2014} \widehat{\beta}_{1}^{(r)}}{\widehat{\theta}_{1}^{(r)}} = 704.$$

- Also, we need parameters estimates for τ_h , ρ_h , and $\chi_{2,h}$ for lags h=11,12. It seems reasonable to set $\widehat{\tau}_h=0$, $\widehat{\rho}_h=0.5$ and $\widehat{\chi}_{2,h}=10,200$ for h=11,12.
- ► The reserve estimate corresponding to claims with longer development is 24, 384, 172.

Comparison with CL

- To enable benchmarking, we include the estimation results as obtained with Chain Ladder (CL) that is the standard reserving technique designed for run-off triangles.
- ► The **results** are the following:

	Reserve estimate	$VaR_{0.95}$	$VaR_{0.995}$
Our approach	27,424,481	29,262,230	30,751,029
CL	22,259,690	24,142,573	25,239,963

Introduction

Approach

Modelling claims with rapid settlement

Modelling claims with longer development

Reserve calculations

References

Short bibliography

- Denuit, M., Charpentier, A. (2005). Mathématiques de l'Assurance Non-Vie. Tome II: Tarification et Provisionnement. Collection Économie et Statistique Avancées, Economica, Paris.
- Denuit, M., Trufin, J. (2016). Collective loss reserving with two types of claims in motor third party liability insurance. Submitted for publication.
- Kaas, R., Goovaerts, M.J., Dhaene, J., Denuit, M. (2008). Modern Actuarial Risk Theory Using R. Springer, New York.
- Pigeon, M., Antonio, K., Denuit, M. (2013). Individual loss reserving with the Multivariate Skew Normal framework. ASTIN Bulletin 43, 399-428.
- Pigeon, M., Antonio, K., Denuit, M. (2014). Individual loss reserving using paid-incurred data. Insurance: Mathematics and Economics 58, 121-131.
- Trufin, J., Albrecher, H., Denuit, M. (2011). Ruin problems under IBNR dynamics. Applied Stochastic Models in Business and Industry 27, 619-632.