Worcester Polytechnic Institute

Safeshop

August, 2013 – Spring 2014

Insurance to Protect Against Online Shopping Fraud

A Major Qualifying Project for A Requirement of the Degree of Bachelor of Science from WORCESTER POLYTECHNIC INSTITUTE

Qiyang Zhou

Major: Actuarial Science
Class of 2014
Advisor: Jon Abraham

Table of Contents

Abstract

This project develops a new type of insurance, online shopping insurance. The project focused on generating a method to calculate the cost of the insurance by collecting data from the Internet Crime Complaint Center's annual reports. I created a simulation model to test the profitability and analyzed ways to reduce the risk. In addition to adding restrictions to the insurance compensation, I introduced the idea of cooperating with an anti-virus software company and analyzed the benefits.

Acknowledgments

I would like to thank Professor Jon Abraham for providing me with guidance and support throughout this project to ensure that I considered all appropriate factors in developing an insurance product. His insight and recommendations were essential to the completion of my project, and I truly appreciate all of his help and support.

Thank you, Professor Abraham,

Qiyang Zhou

Basic ideas

What is it?

Nowadays more and more people are shopping online, but it is not always safe to do that. Sometimes people get tricked by sellers, and they may never get the item they paid for. When that happens, people may face an unpleasant loss. The basic idea of this insurance is to cover policyholders' shopping loss caused by Internet fraud.

Why do people need it?

Imagine someone paid about $800 for a computer online, but he never received this computer. Then he wasted $800, and he could report this issue to www.ic3.gov (The Internet Crime Complaint Center), a partner with the FBI and the National White Collar Crime Center. He will probably report this case through a long process, and he might need to go to court in the future to pursue compensation from the seller. This will cost him a lot of time and money. Instead, he might want to try online shopping insurance if the insurance company offers him a good price, such as $10. If he does not get the computer, the insurer pays him $800 and the insurance company will get the money back from the seller through the legal process.

Do a lot of people need this insurance?

According to the census stats from 2012, the population in the United States is 313,847,465, and 245,203,319 (78.1%) of them are using the Internet. Eighty-one percent of Internet users research products online. Twenty-six percent of Internet users have used online auctions. Sixty-six percent of Internet users have purchased an online product. Thus, numerous people are shopping online. From the annual report of www.ic3.gov for 2012, 289,874 suffered losses from online shopping, and the total loss was $525,441,110.00. Since there are a lot of people who suffered losses from online shopping, people would like to buy this insurance.

Demand Analysis

Growth of E-commerce

Michael Aldrich invented online shopping in 1979, and Tim Berners-Lee created the first World Wide Web server and browser in 1990. Later, in 1994 and 1995, Netscape, Pizza Hut, Amazon, and EBay started to sell their products online. For recent years, E-commerce sales are growing at 10% each year in the United States and 19% worldwide.

Figure 1

(source: http://www.groupin.pk/blog/e-commerce-facts-current-trends-future-prospects/)

Figure 2

(source: http://www.groupin.pk/blog/e-commerce-facts-current-trends-future-prospects/)

E-commerce has been changing people's lives

E-commerce started in the mid-1990s, and it has been changing the business model in the world. From the 2012 report of WorldPay, most people would like to spend more than 1/6 of their disposable income on online shopping and people worldwide would like to spend 22% of their disposable income on online shopping, on average.

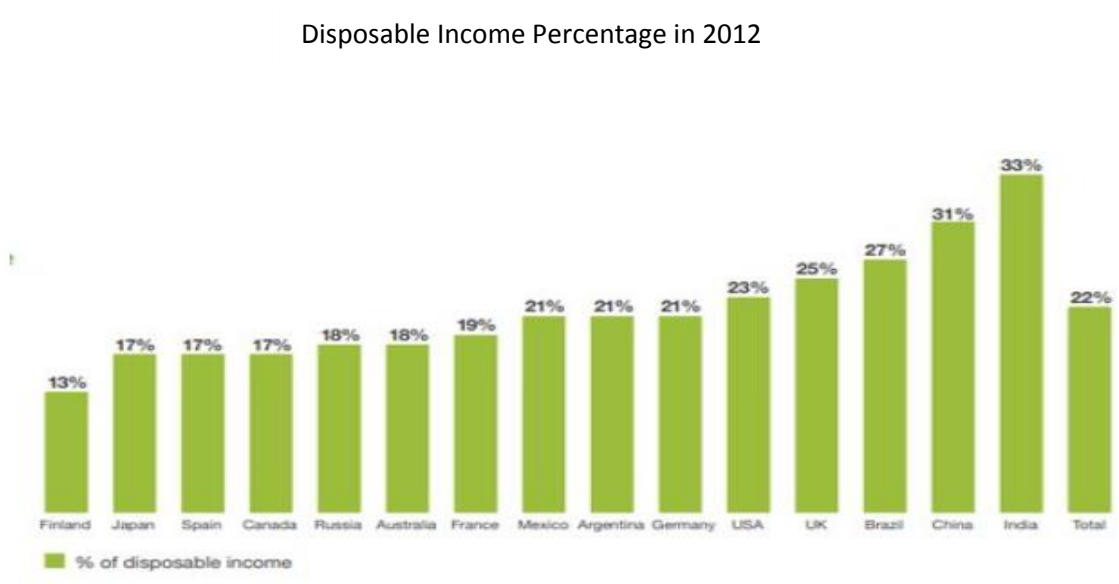


Figure 3

(Data from http://www.worldpay.com/about_us/content/worldpay-annual-report-2012.pdf.)

The shopping hot time has changed from daytime to nighttime. Worldwide, 74% of online transactions took place between noon and midnight, and most of them happened during the nighttime. Retailers worldwide soon noticed that they need to adapt their business model to the new situation that people like to shop more often at night.

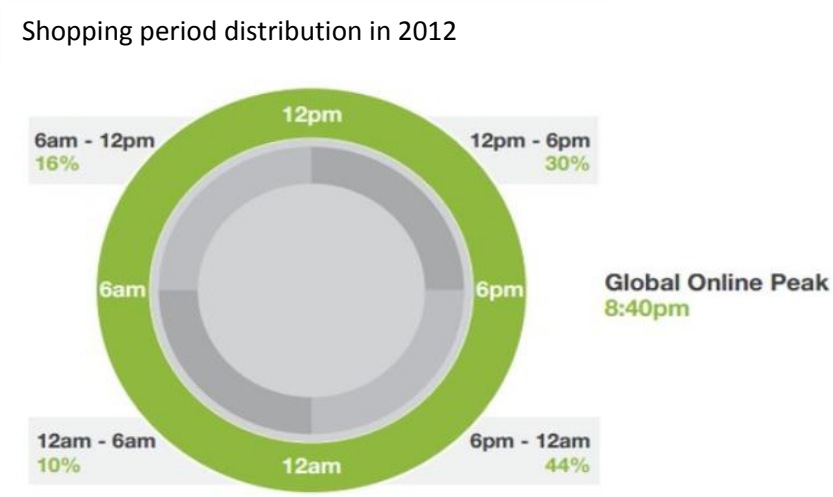


Figure 4

(Data from http://www.worldpay.com/about_us/content/worldpay-annual-report-2012.pdf.)

Risks of E-commerce

Like a coin has two sides, e-commerce brings convenience but also risks to customers. In recent years, the rapid development of e-commerce has increased online transactions but also online fraud. More and more people get tricked by online "sellers" in various ways. From the 2012 annual report of www.ic3.com, 289,874 people lost money by shopping online, and the total loss was $525,441,110.00. The average dollar loss overall was $1813.00 and the average dollar loss for those reporting a loss was $4,573.00.

Figure 5

(Data collected from www.ic3.gov)

Figure 6

(Data collected from www.ic3.gov)

Why should we have E-commerce insurance (Safeshop)?

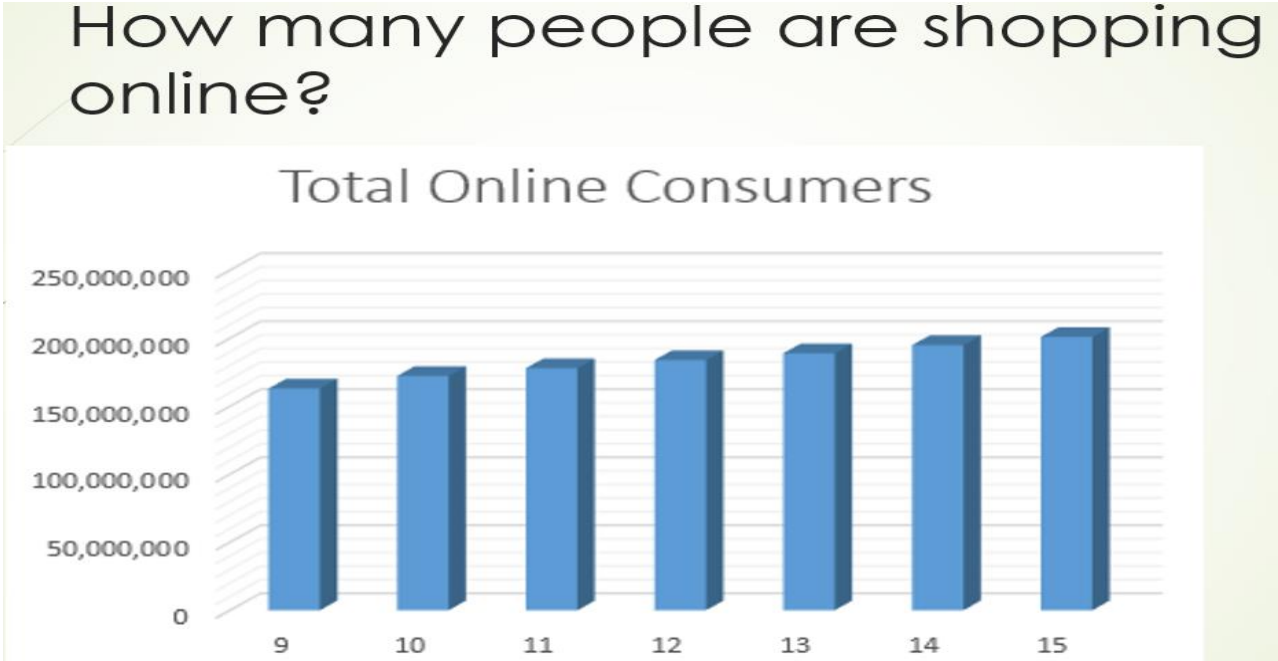


Figure 7

(Data collected from http://www.groupin.pk/blog/e-commerce-facts-current-trends-future-prospects/)

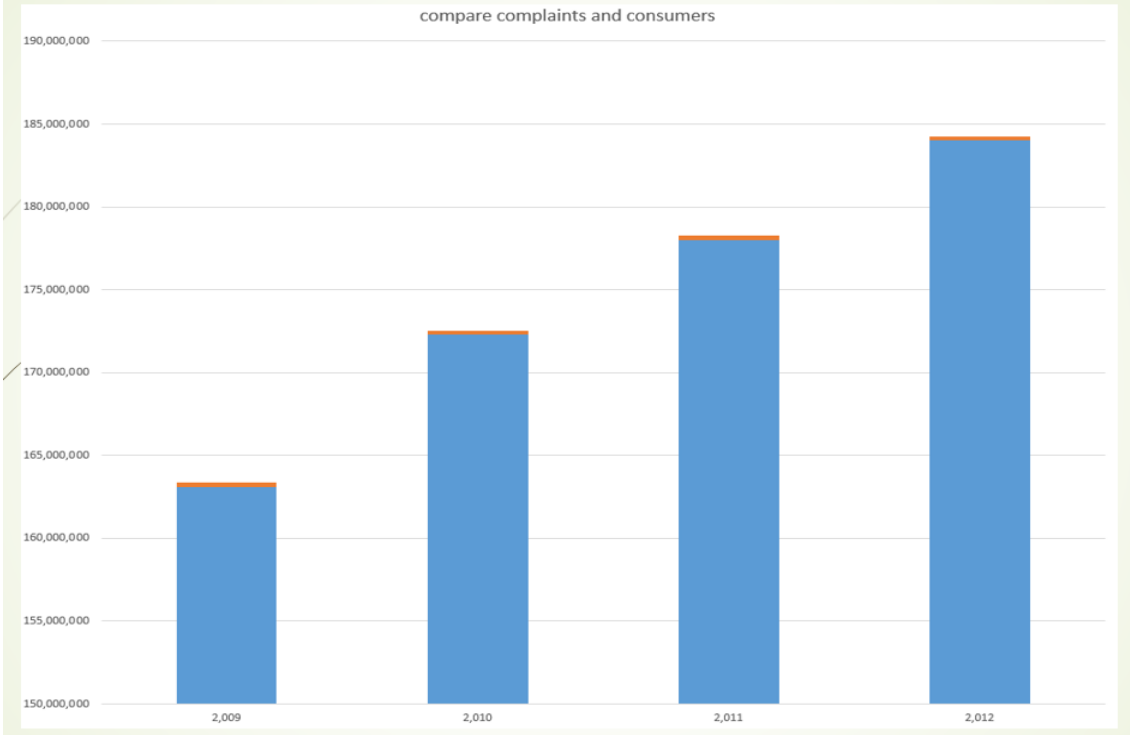


Figure 8

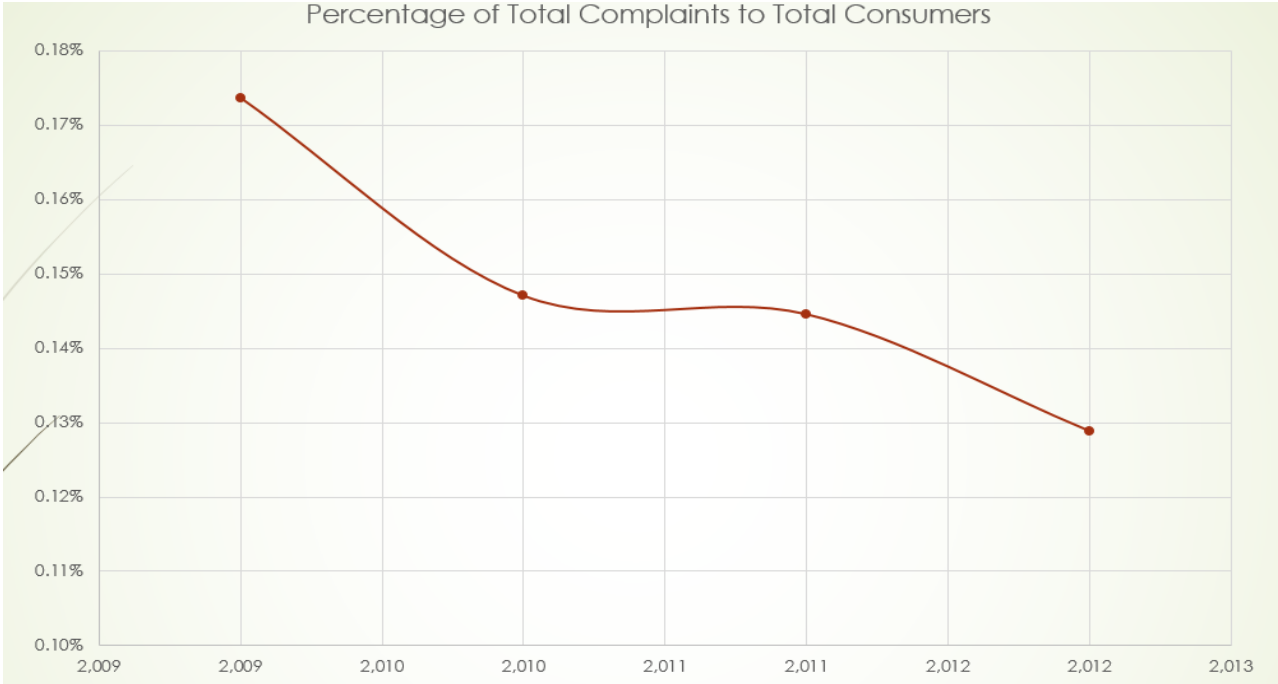


Figure 9

From the figures above, it is obvious that the ratio of people who have online shopping losses to total online consumers is not high. The rate is about 0.16%, thus for every 100,000 online consumers, there will only be about 160 people who will have a loss related to online fraud throughout a year.

From the 2012 Household Income report of www.census.gov, the average annual income for the US Median Household was $51,371. Thus, one loss from online shopping would probably cost the household a month's salary, which means they have worked one month for nothing. People are not willing to take this loss. Furthermore, in order to get back the money loss, complainants would report this issue to www.ic3.gov, and they might spend more time and money to go through the legal process. For the worst situation, they might not get back any money loss after all the legal process. So, the online loss happens at a low rate, but it will cost people a lot when it happens. Today, there is no such insurance to protect the loss of online shoppers; therefore there are complaints and the government has to deal with a large amount of aftermath and cost, and the resulting amount of compensation for victims is also very limited. In accordance with the rules of a market economy, if we can play a role as insurance, it will benefit both online shoppers and the government. As a result, the online shopping insurance could

reduce the financial burden of the government, transfer the loss risk of the online shoppers, and protect the legitimate interests of the victims.

Cost of the Insurance

Build Simple Models

In order to find out the method for calculating the cost of the insurance, and set up proper restrictions of the insurance in a short amount of time, I decided to simulate the loss models with Matlab by using simple data first.

Build models to set up price for a 1000-consumer insurance plan

According to the IC3's annual report and the e-commerce annual report from the government census website (www.census.gov), there are about 160 people who lost money by shopping online for every 100,000 people each year. I used Matlab to model this situation. First, I built a data set that contains 100,000 elements; 160 of them lost money and others (99840) did not. Thus, 160 elements (loss) would be values that were greater than 0, and the rest (99840 elements) would appear to be 0. Second, I gave values (a1 a2 a3 ... a160) to non-zero elements (160 losses) and fitted them into a distribution. Therefore, the data would look like {a1 a2 a3 ... a160 0 0 0 0 .....}.

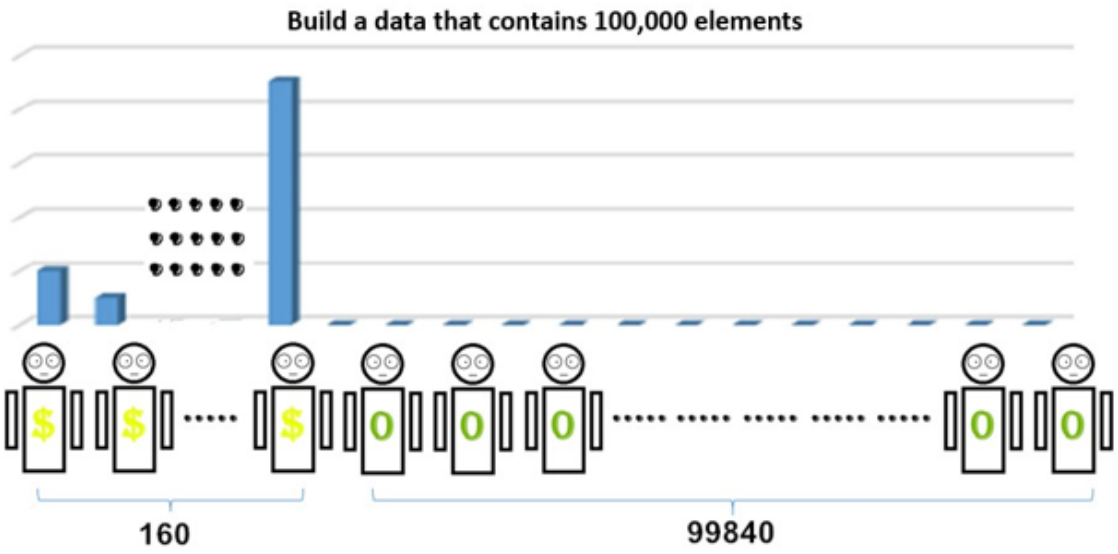


Figure 10

Next, I would repeat the experiment, and each time I took 1,000 elements from the sample space (100,000). If any of them reports a loss (non-zero elements), this insurance will cover it. For each experiment, I supposed the insurance price for each person to be $5; thus the

total sale would be $5,000 (5*1,000). This insurance would pay policyholders a total amount of money equal to the sum of the non-zero elements. For example, if one experiment selects 1,000 elements as {a1 a3 a7 a100 a125 0 0 0 0 0 ... 0 0 0} (5 non-zero elements and 995 zero elements), the total sale will be $5,000, the total insurance compensation will be $(a1+a3+a7+a100+a125), and the profit will be $5,000 - $(a1+a3+a7+a100+a125).

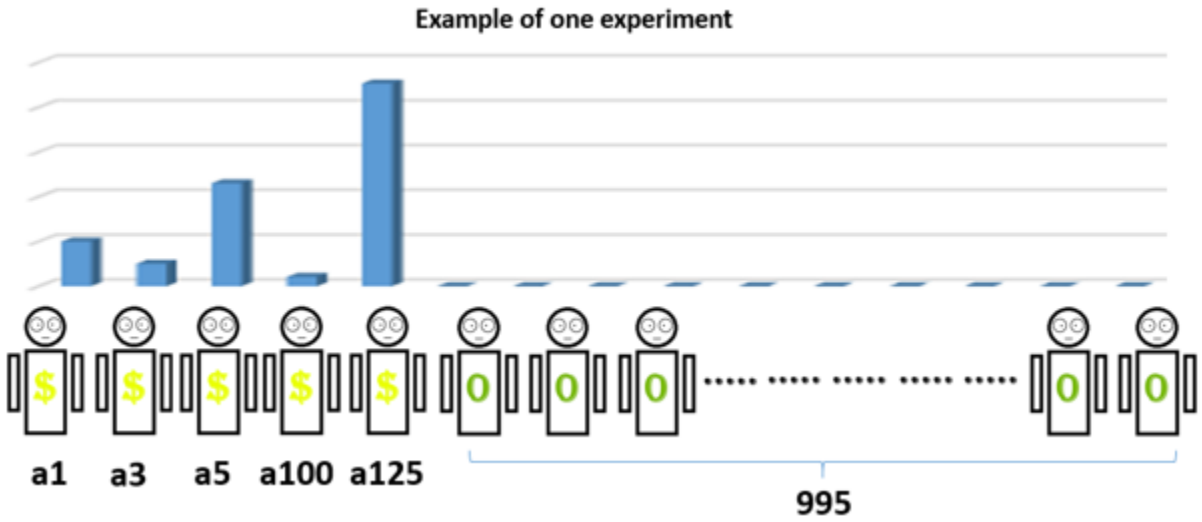


Figure 11

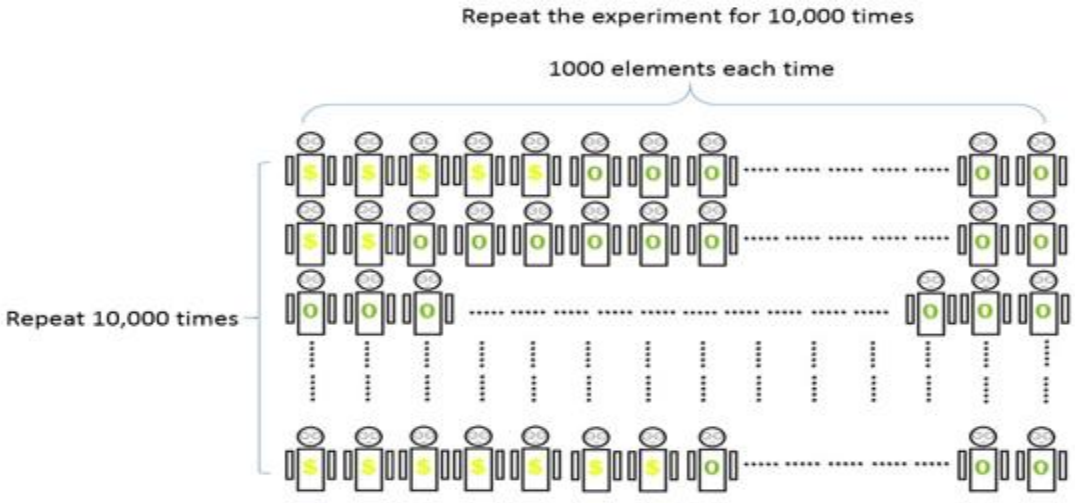


Figure 12

I recorded the insurance (sum of the non-zero elements) and profit (5,000 - insurance) for each experiment, and the insurance was the money that the insurance program was going to send out to the policyholders. After repeating the experiment 10,000 times, there were 10,000 insurances. Then, I compared the 10,000 insurances, and calculated the minimum (min), the maximum (max), the average (total/10,000), and the expected value of those experiments.

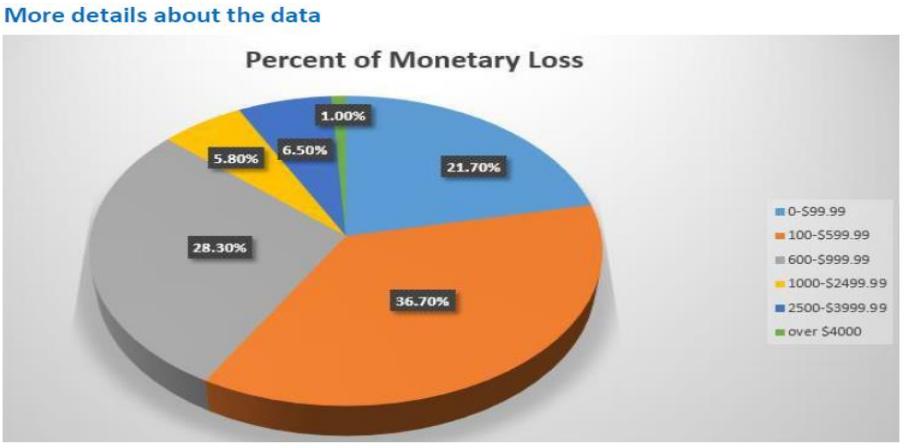


Figure 13

(data from ic3's annual reports from 2006 to 2012)

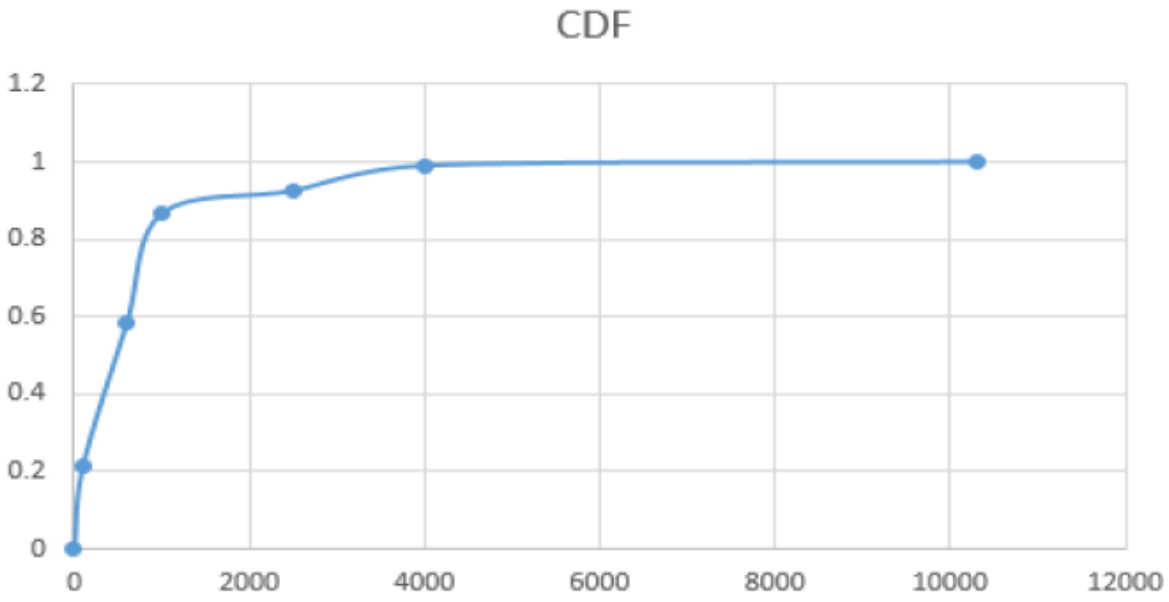


Figure 14

From those graphs above, I found that the overall data may fit a lognormal distribution or gamma distribution, and then I used inverse transform method to simulate the data.

For 10,000 times

```
Summary =
                          Min           Max       Average        Expect
     Insurance              0   26571.08092    1353.61512    1343.02322
         Price              0      26.57108       1.35362       1.34302
        Profit -21571.08092    5000.00000    3646.38488    3656.97678


gainamount =

         9626


gainpercent =

     0.9626
```

Figure 15

As the summary showed from this experiment, the minimum insurance was 0 (1,000 elements selected were all policyholders with zero loss), the maximum insurance was about $26,571 (the sum of the non-zero elements was $26,571 for one experiment), and the average insurance was about $1,354 (the sum of the 10,000 insurance divided by 10,000) which was close to the expected value, $1,343.

Since the total sale was $5,000, the minimum profit would be

5,000 – (max insurance) = 5,000 – 26,571 = -26,571

the maximum profit would be

5,000 – (min insurance) = 5,000

and the average profit would be

5,000 – (average insurance) = 5,000 – 1,354 = 3,646

The price showed on the summary above was the price to make the profit to be 0 (cost of the insurance).

Thus, the maximum price would be

(max insurance)/1,000 = 26,571/1,000 = 26.571

the minimum price would be

(min insurance)/1,000 = 0,

and the average price would be

(average insurance)/1,000 = 1,354/1,000 = 1.354.

The "gainamount" shown on the summary above was the number of times when the profit was not below zero, and the "gainpercent" was "gainamount"/100,000 (profitable experiments/total experiments). Therefore, 9,626 experiments among 10,000 experiments were not losing money if the insurance price was $5 per policyholder.

We can see that the actual value gets closer to the expected value as I do more iterations. I am concerned that there is a chance to lose more than 20,000 dollars from this insurance, which is four times more than the max benefits we can get. To make sure the insurance company does not lose such a large amount of money from this insurance, it has to set up a price more than $26, which may not interest the buyers. The insurance company would not want to take this risk, so I set up restrictions and examine how they will affect the cost and profit.

Adjust the insurance system

In order to reduce the maximum loss the insurance is going to face, I was trying to add restrictions to the pricing system such as maximum payment, deductibles, and percentage payment.

For 10,000 times (maximum payment)

Summary =

	Min-Profit	Average-Prof	Max-Insuranc	Max-Price	Gain-Percent
500	2204.02213	4471.28682	2795.97787	2.79598	1.00000
1000	608.90181	4225.21872	4391.09819	4.39110	1.00000
1500	-1151.46421	4125.67581	6151.46421	6.15146	0.99910
2000	-2607.12132	4050.15952	7607.12132	7.60712	0.99710
2500	-3607.12132	3990.38178	8607.12132	8.60712	0.99060
3000	-4666.18022	3942.08346	9666.18022	9.66618	0.98700
3500	-6166.18022	3906.10780	11166.18022	11.16618	0.98060
4000	-7037.51337	3882.64879	12037.51337	12.03751	0.97540
no-max	-20353.16481	3722.36123	25353.16481	25.35316	0.96210

Figure 16

safe price	max payment of a year
2.80	500
4.40	1,000
6.16	1,500
7.61	2,000
8.60	2,500
9.67	3,000
11.17	3,500
12.04	4,000
25.36	no max

Figure 17

The safe price is the price (regardless of expenses) to make the insurance plan have no loss in all of the simulated situations. As a result, the average profit becomes larger than before, and if we set up the maximum payment to be $500, we are very likely to earn money. If we set up the price to be $7 a year, we can have the premium be $1500 and still gain a profit. If we set the premium to be larger than $4000, we have to set up the price to be larger than $15 in order to reduce the risk of losing an unpleasant amount of money.

For 10,000 times (deductible)

Assign numbers to uniform distribution with the same sample data each time as the deductible goes up, and the insurance only covers the loss over 100. For example if the deductible is 100, and the loss is 500, the insurance covers 400 of the loss.

Summary =

	Min-Profit	Average-Prof	Max-Insuranc	Max-Price	Gain-Percent
no-min	-20353.16481	3722.36123	25353.16481	25.35316	0.96210
50	-20203.16481	3797.34089	25203.16481	25.20316	0.96540
100	-20053.16481	3864.47543	25053.16481	25.05316	0.96770
150	-19903.16481	3924.54268	24903.16481	24.90316	0.96950
200	-19753.16481	3980.91337	24753.16481	24.75316	0.97140
250	-19603.16481	4033.04093	24603.16481	24.60316	0.97280
300	-19453.16481	4082.41964	24453.16481	24.45316	0.97390
350	-19303.16481	4129.47745	24303.16481	24.30316	0.97460
400	-19153.16481	4173.04610	24153.16481	24.15316	0.97550

safe price	deductible
25.36	no
25.21	50
25.06	100
24.91	150
24.76	200
24.61	250
24.46	300
24.31	350
24.16	400

Figure 18 and figure 19

For 10,000 times (percentage compensation)

Assign number to uniform distribution with same sample data each time, and the insurance only covers a certain percent of the total loss. For example if the percentage is 75%, and the loss is 100, the insurance covers 75, and the policyholder will pay 25 by himself.

Summary =

	Min-Profit	Average-Prof	Max-Insuranc	Max-Price	Gain-Percent
100%	-20353.16481	3722.36123	25353.16481	25.35316	0.96210
95%	-19085.50657	3786.24317	24085.50657	24.08551	0.96650
90%	-17817.84833	3850.12511	22817.84833	22.81785	0.97000
85%	-16550.19009	3914.00704	21550.19009	21.55019	0.97320
80%	-15282.53185	3977.88898	20282.53185	20.28253	0.97490
75%	-14014.87361	4041.77092	19014.87361	19.01487	0.97650
70%	-12747.21537	4105.65286	17747.21537	17.74722	0.97750
65%	-11479.55713	4169.53480	16479.55713	16.47956	0.97820
60%	-10211.89889	4233.41674	15211.89889	15.21190	0.97840

safe price	percentage
25.36	100%
24.09	95%
22.82	90%
21.56	85%
20.28	80%
19.02	75%
17.75	70%
16.48	65%
15.22	60%

Figure 20 and figure 21

Combination of maximum payment and deductibles

Assign number to uniform distribution with same sample data each time. The insurance covers the loss beyond deductibles and there is a maximum of the insurance. For example, if the deductible is 100 and maximum payment is 1000, the insurance covers the loss beyond 100 and the insurance will not exceed 1000.

safe price	500	1000	1500	2000	2500	3000	3500	4000	no-max
400	2.10	4.00	5.66	6.66	7.77	9.27	10.44	10.44	24.16
350	2.22	4.00	5.72	6.76	7.82	9.32	10.64	10.64	24.31
300	2.32	4.00	5.77	6.86	7.87	9.37	10.79	10.84	24.46
250	2.37	4.00	5.82	6.96	7.96	9.42	10.89	11.04	24.61
200	2.42	4.00	5.87	7.06	8.06	9.47	10.97	11.24	24.76
150	2.47	4.00	5.92	7.16	8.16	9.52	11.02	11.44	24.91
100	2.51	4.05	6.02	7.31	8.31	9.57	11.07	11.64	25.06
50	2.65	4.18	6.10	7.46	8.46	9.62	11.12	11.84	25.21
no deductible	2.80	4.39	6.15	7.61	8.61	9.67	11.17	12.04	25.36

Figure 22

Combination of maximum payment and percentages

Assign number to uniform distribution with same sample data each time. The insurance covers the loss for a certain percentage and there is a maximum of the insurance. For example, if the percentage is 80% and maximum payment is 1000, the insurance covers 80% of the loss and the insurance will not exceed 1000.

safe price	500	1000	1500	2000	2500	3000	3500	4000	no-max
60%	2.42	4.03	5.16	6.40	7.22	7.22	7.98	8.98	15.22
65%	2.51	4.10	5.34	6.43	7.82	7.82	8.07	9.07	16.48
70%	2.56	4.11	5.52	6.52	7.97	8.43	8.43	9.15	17.75
75%	2.60	4.11	5.71	6.71	8.00	9.03	9.03	9.23	19.02
80%	2.64	4.12	5.88	6.89	8.03	9.51	9.63	9.63	20.28
85%	2.68	4.13	5.96	7.07	8.07	9.57	10.23	10.23	21.56
90%	2.72	4.15	6.05	7.25	8.25	9.60	10.83	10.83	22.82
95%	2.76	4.27	6.13	7.43	8.43	9.63	11.13	11.44	24.09
100%	2.80	4.39	6.15	7.61	8.61	9.67	11.17	12.04	25.36

Figure 23

Combination of percentages and deductibles

Assign number to uniform distribution with same sample data each time. The insurance covers a certain percent of loss beyond the deductibles. For example, if the percentage is 80% and the deductible is 100, the insurance covers 80% of the loss beyond 100.

safe price	400	350	300	250	200	150	100	50	none
60%	14.49	14.58	14.67	14.76	14.85	14.94	15.03	15.12	15.21
65%	15.70	15.80	15.89	15.99	16.09	16.19	16.28	16.38	16.48
70%	16.91	17.01	17.12	17.22	17.33	17.43	17.54	17.64	17.75
75%	18.11	18.23	18.34	18.45	18.56	18.68	18.79	18.90	19.01
80%	19.32	19.44	19.56	19.68	19.80	19.92	20.04	20.16	20.28
85%	20.53	20.66	20.79	20.91	21.04	21.17	21.30	21.42	21.55
90%	21.74	21.87	22.01	22.14	22.28	22.41	22.55	22.68	22.82
95%	22.95	23.09	23.23	23.37	23.52	23.66	23.80	23.94	24.09
100%	24.15	24.30	24.45	24.60	24.75	24.90	25.05	25.20	25.35

Figure 24

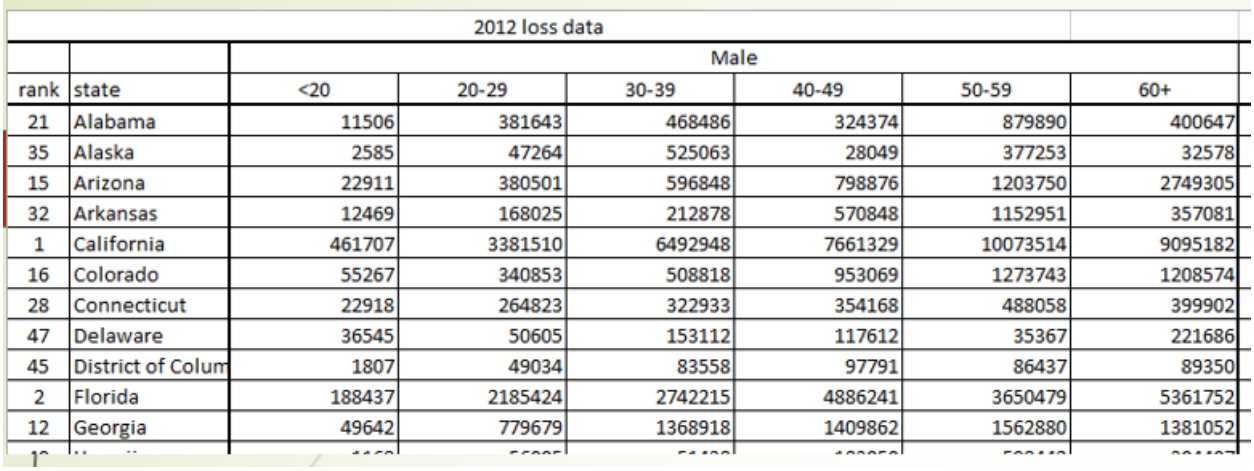
Conclusion

From all of the simulations, I found out that the max payment would be the best insurance strategy to help reduce the risk and keep the insurance product competitive.

Detailed analysis of the cost of the insurance

Collect the data

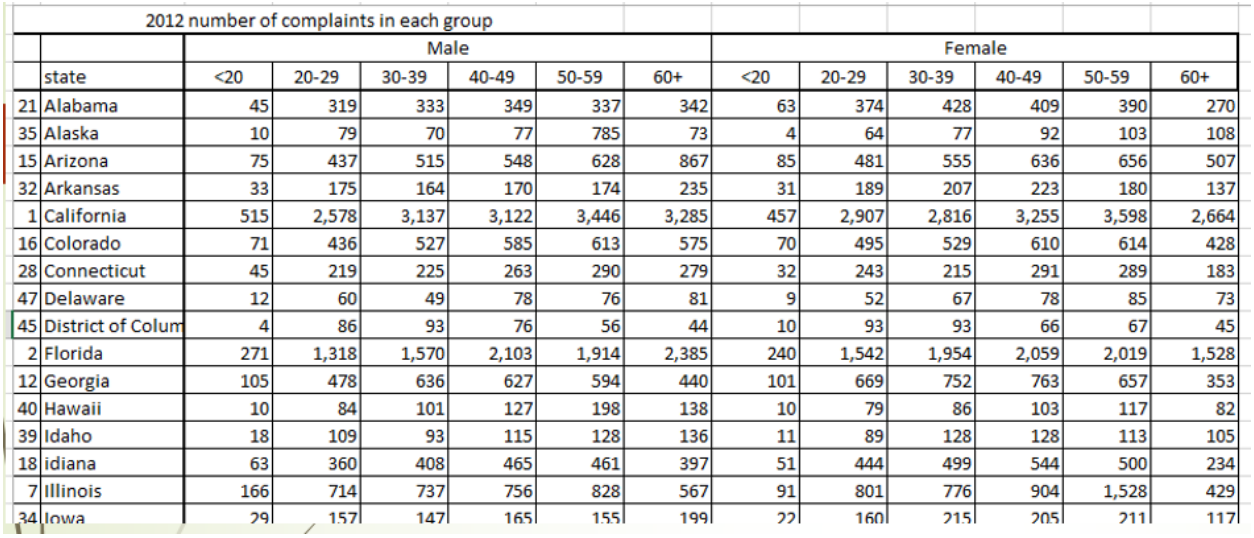
I looked through annual reports from 2006 to 2012 on the government website, www.ic3.gov. For the data (total loss and total complaints) in each year, I divided them into 612 different groups by their states (51 states including the District of Columbia), gender (male, female), and age (groups, 6 different groups). Below is part of the loss data in 2012.



2012 loss data							
		Male					
rank	state	<20	20-29	30-39	40-49	50-59	60+
21	Alabama	11506	381643	468486	324374	879890	400647
35	Alaska	2585	47264	525063	28049	377253	32578
15	Arizona	22911	380501	596848	798876	1203750	2749305
32	Arkansas	12469	168025	212878	570848	1152951	357081
1	California	461707	3381510	6492948	7661329	10073514	9095182
16	Colorado	55267	340853	508818	953069	1273743	1208574
28	Connecticut	22918	264823	322933	354168	488058	399902
47	Delaware	36545	50605	153112	117612	35367	221686
45	District of Colum	1807	49034	83558	97791	86437	89350
2	Florida	188437	2185424	2742215	4886241	3650479	5361752
12	Georgia	49642	779679	1368918	1409862	1562880	1381052

Figure 25

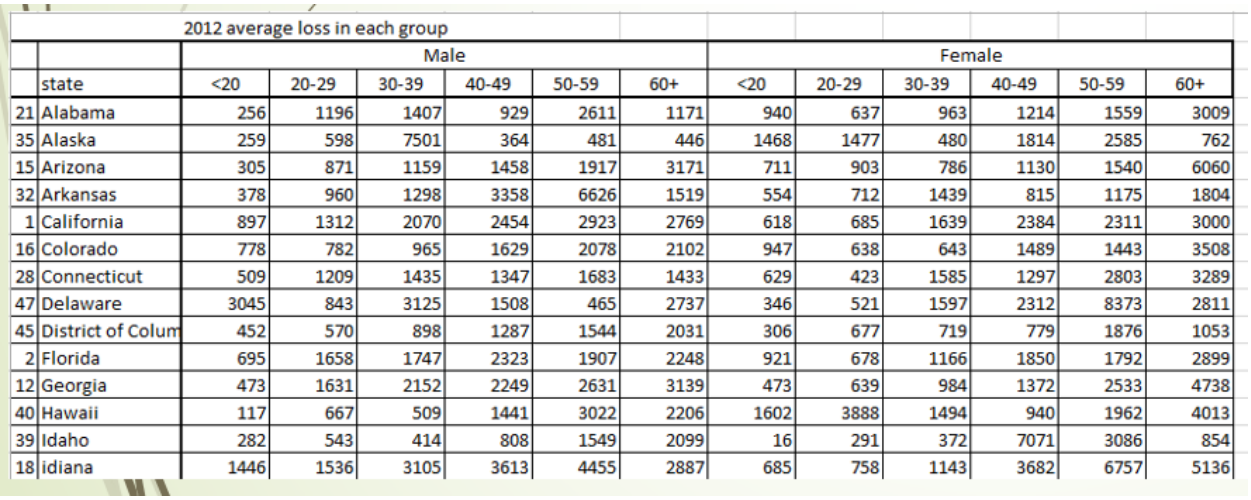
The data above shows the total money loss in each group.



2012 number of complaints in each group													
		Male						Female					
	state	<20	20-29	30-39	40-49	50-59	60+	<20	20-29	30-39	40-49	50-59	60+
21	Alabama	45	319	333	349	337	342	63	374	428	409	390	270
35	Alaska	10	79	70	77	785	73	4	64	77	92	103	108
15	Arizona	75	437	515	548	628	867	85	481	555	636	656	507
32	Arkansas	33	175	164	170	174	235	31	189	207	223	180	137
1	California	515	2,578	3,137	3,122	3,446	3,285	457	2,907	2,816	3,255	3,598	2,664
16	Colorado	71	436	527	585	613	575	70	495	529	610	614	428
28	Connecticut	45	219	225	263	290	279	32	243	215	291	289	183
47	Delaware	12	60	49	78	76	81	9	52	67	78	85	73
45	District of Colum	4	86	93	76	56	44	10	93	93	66	67	45
2	Florida	271	1,318	1,570	2,103	1,914	2,385	240	1,542	1,954	2,059	2,019	1,528
12	Georgia	105	478	636	627	594	440	101	669	752	763	657	353
40	Hawaii	10	84	101	127	198	138	10	79	86	103	117	82
39	Idaho	18	109	93	115	128	136	11	89	128	128	113	105
18	idiana	63	360	408	465	461	397	51	444	499	544	500	234
7	Illinois	166	714	737	756	828	567	91	801	776	904	1,528	429
34	Iowa	29	157	147	165	155	199	22	160	215	205	211	117

Figure 26

The data above shows the total complaints in each group.



2012 average loss in each group												
		Male						Female					
	state	<20	20-29	30-39	40-49	50-59	60+	<20	20-29	30-39	40-49	50-59	60+
21	Alabama	256	1196	1407	929	2611	1171	940	637	963	1214	1559	3009
35	Alaska	259	598	7501	364	481	446	1468	1477	480	1814	2585	762
15	Arizona	305	871	1159	1458	1917	3171	711	903	786	1130	1540	6060
32	Arkansas	378	960	1298	3358	6626	1519	554	712	1439	815	1175	1804
1	California	897	1312	2070	2454	2923	2769	618	685	1639	2384	2311	3000
16	Colorado	778	782	965	1629	2078	2102	947	638	643	1489	1443	3508
28	Connecticut	509	1209	1435	1347	1683	1433	629	423	1585	1297	2803	3289
47	Delaware	3045	843	3125	1508	465	2737	346	521	1597	2312	8373	2811
45	District of Colum	452	570	898	1287	1544	2031	306	677	719	779	1876	1053
2	Florida	695	1658	1747	2323	1907	2248	921	678	1166	1850	1792	2899
12	Georgia	473	1631	2152	2249	2631	3139	473	639	984	1372	2533	4738
40	Hawaii	117	667	509	1441	3022	2206	1602	3888	1494	940	1962	4013
39	Idaho	282	543	414	808	1549	2099	16	291	372	7071	3086	854
18	idiana	1446	1536	3105	3613	4455	2887	685	758	1143	3682	6757	5136

Figure 27

The data above shows the average loss in each group.

Predict the future data

In order to set up the loss model, I need to predict the data trend in 2014. So I need to use the historical data I have (from 2006 to 2012) to predict the data in 2014. Then I look at the data trend in each group from 2006 to 2012.

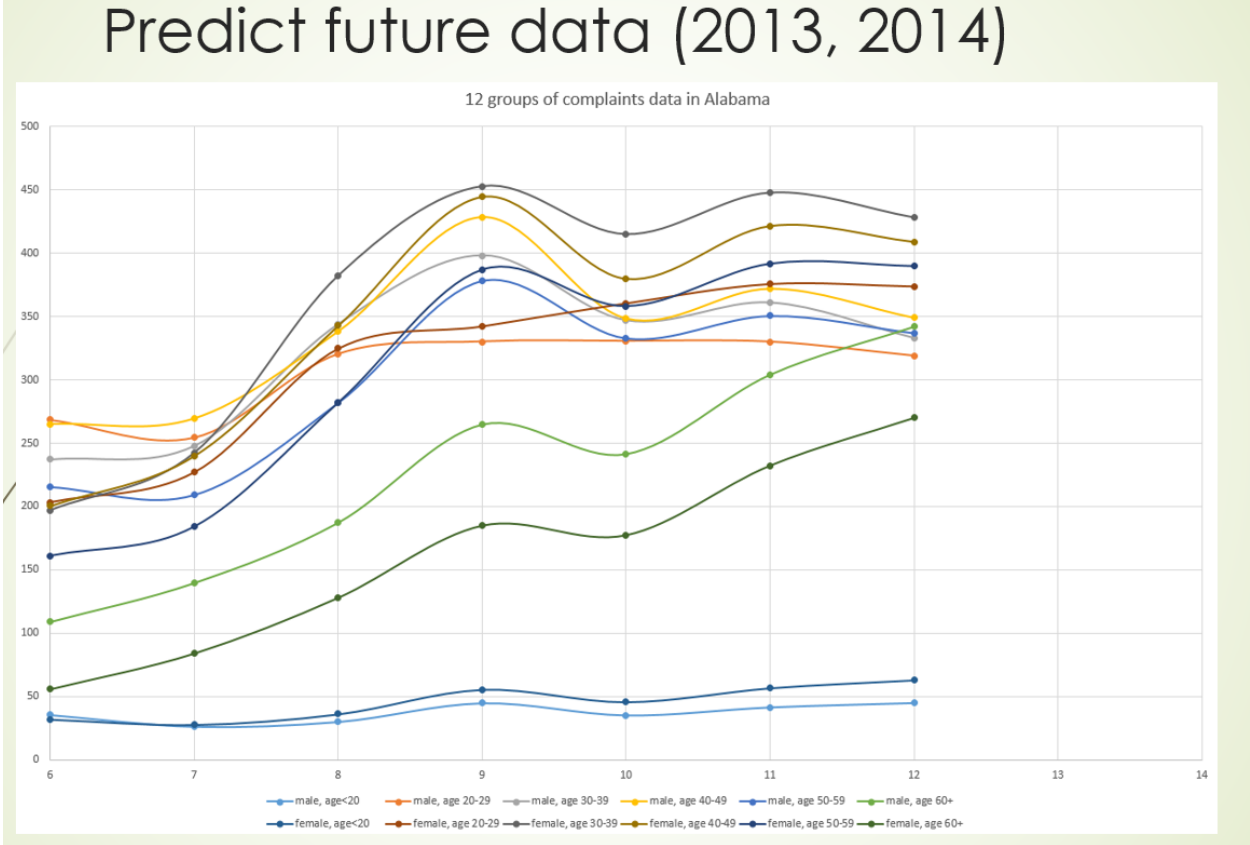


Figure 28

Using the 12 groups of data (total complaints) in Alabama as an example, I found out that the trend goes linear except for the data in 2009. I thought that the economic crisis in 2009 might be a reason for this difference. I looked at the GDP growth rate and it supported my thought.

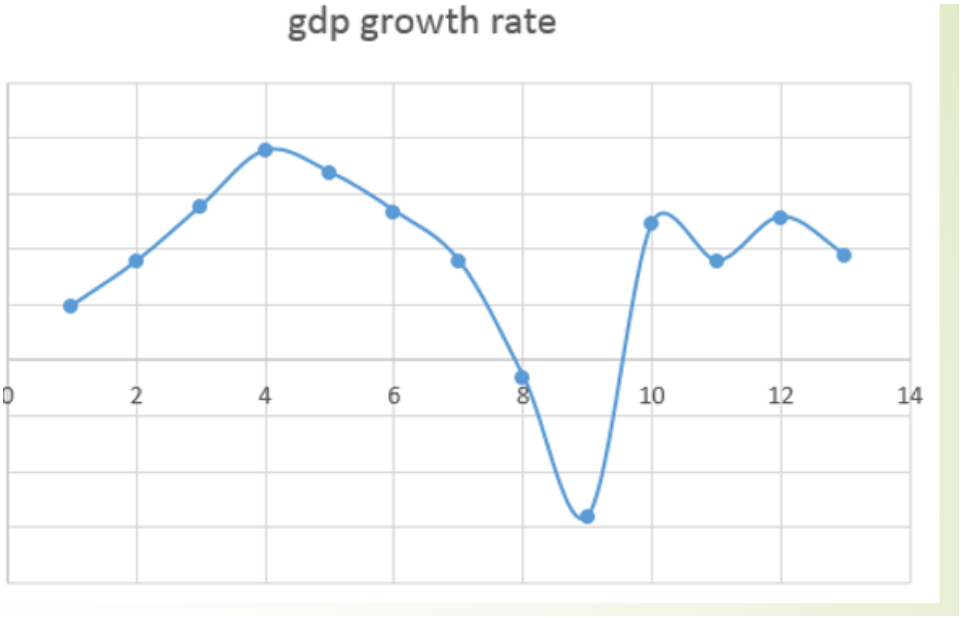


Figure 29

Therefore, I decided not to use the data in 2009 to get a better prediction.

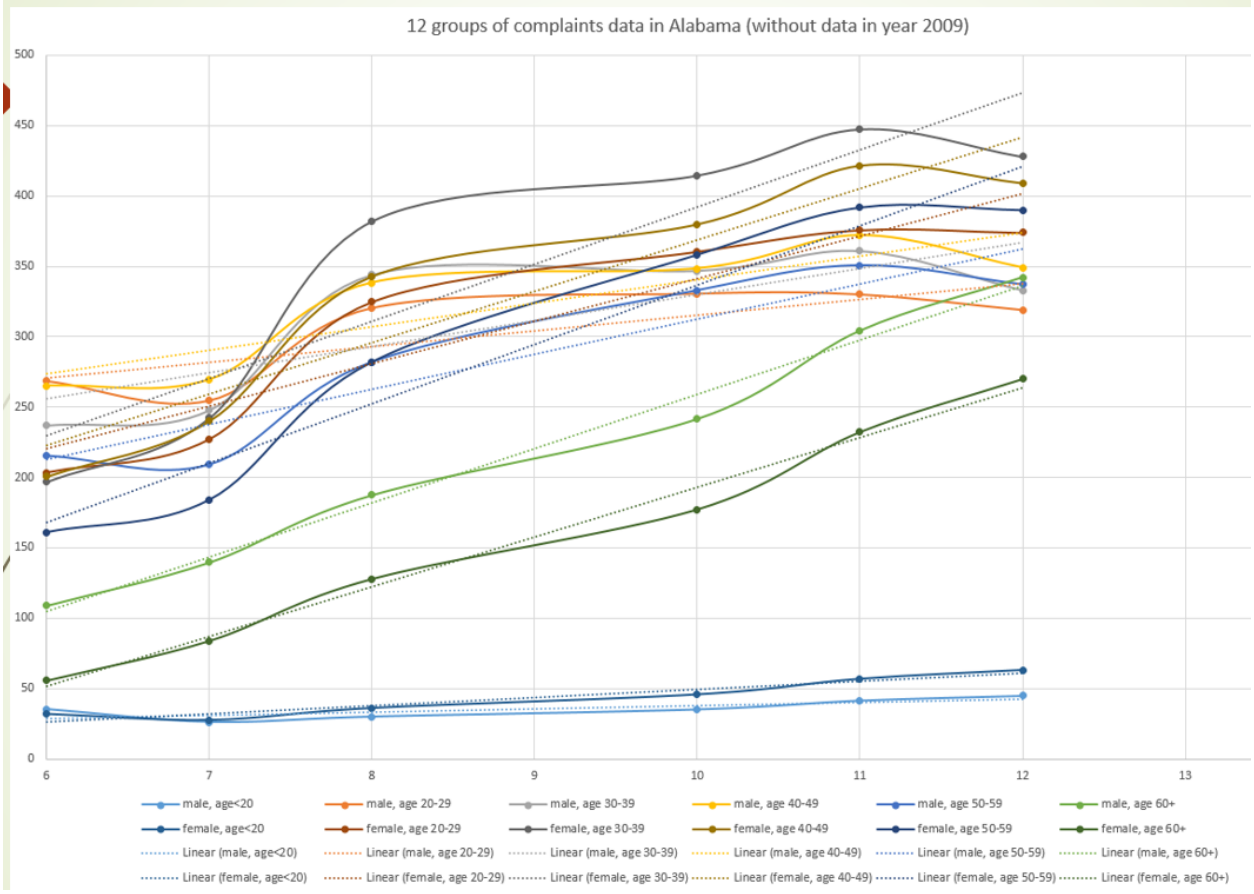


Figure 30

As most of the trends are linear, I decided to use linear least square to predict the data in 2013 and 2014.

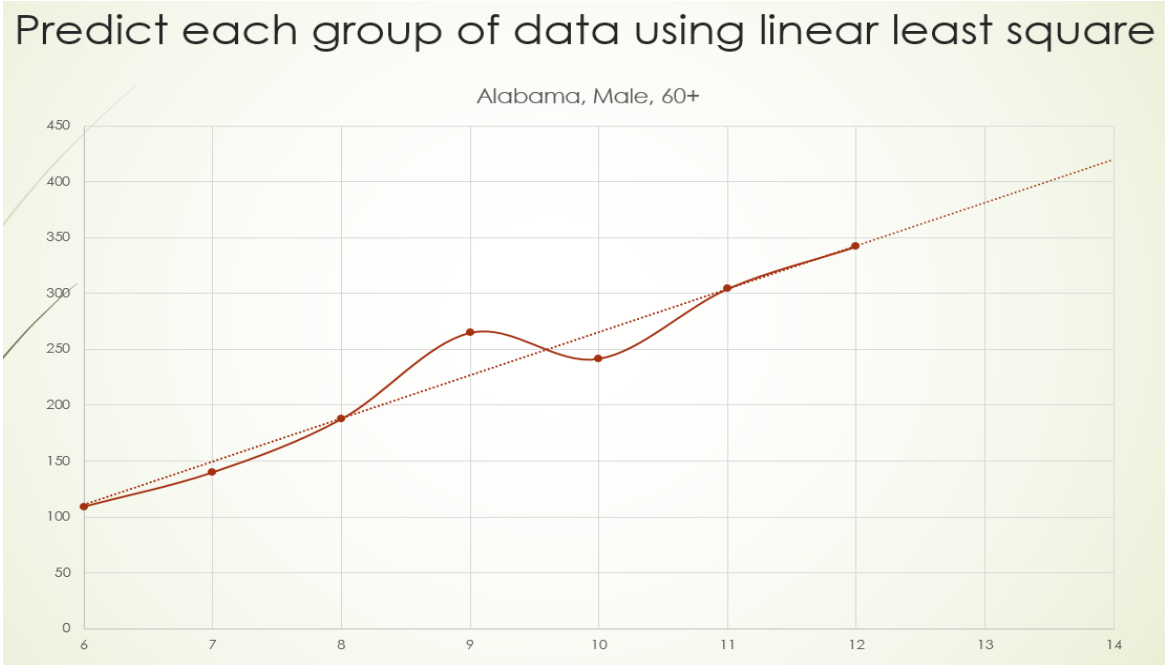


Figure 31

The figure above is an example of how I predicted the future data in each group. Then I have the predicted data (total complaints in each group) for 2014. When I used the same strategy to predict the loss data (total money loss in each group) for 2014, I multiplied each loss data by their relative value of a dollar. For example, $1 in 2006 will equal $1.21 in 2014, so I multiplied the loss data in 2006 by 1.21.

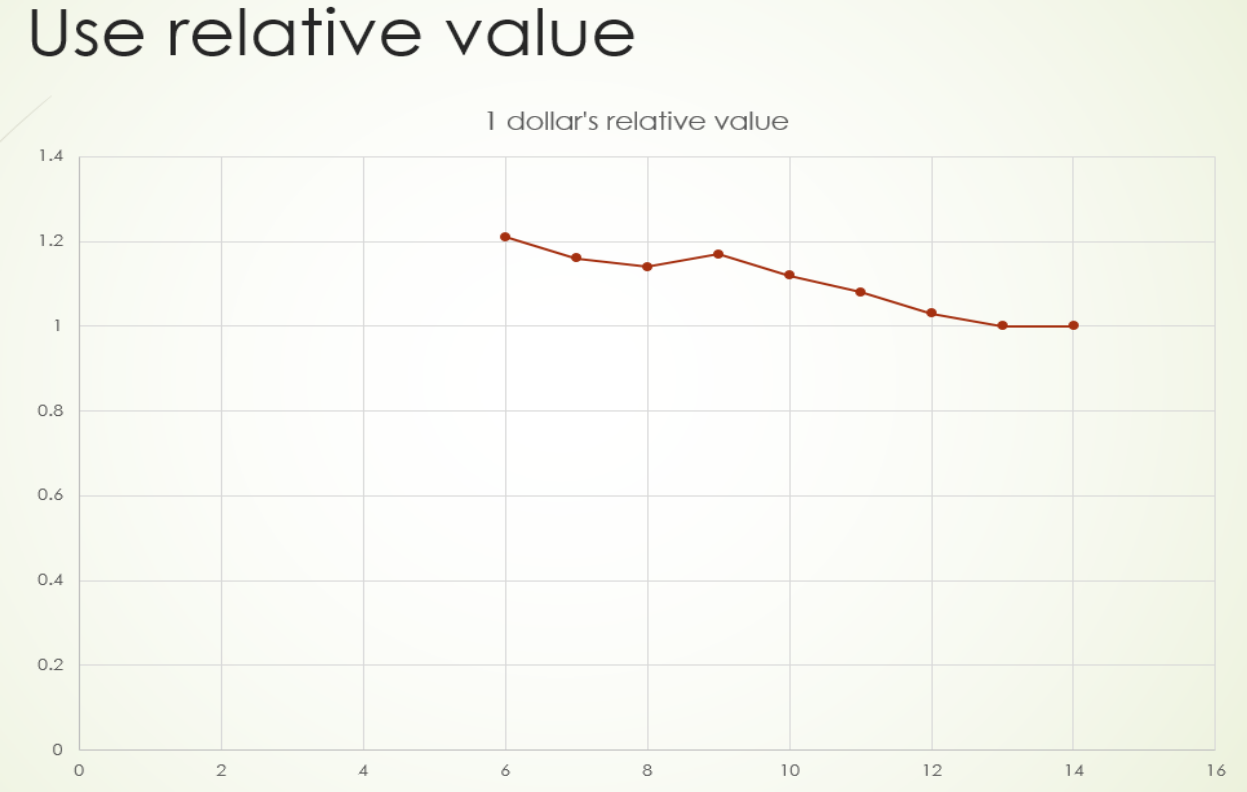


Figure 32

(Data calculated from http://www.measuringworth.com/)

Then I gained the total complaints data and total loss data for each group in 2014. In order to get the loss model, I still needed the loss distribution of each group. Since I divided each group by similar characteristics (same states, same gender, and same age groups), I assumed that each of them would form a normal distribution with corresponding parameters (mean and standard deviation). Then, I used the historical data (from 2006 to 2012) to calculate the mean and standard deviation for each group.

Using one group (Alabama, male, under 20 years old) as an example to show the calculations:

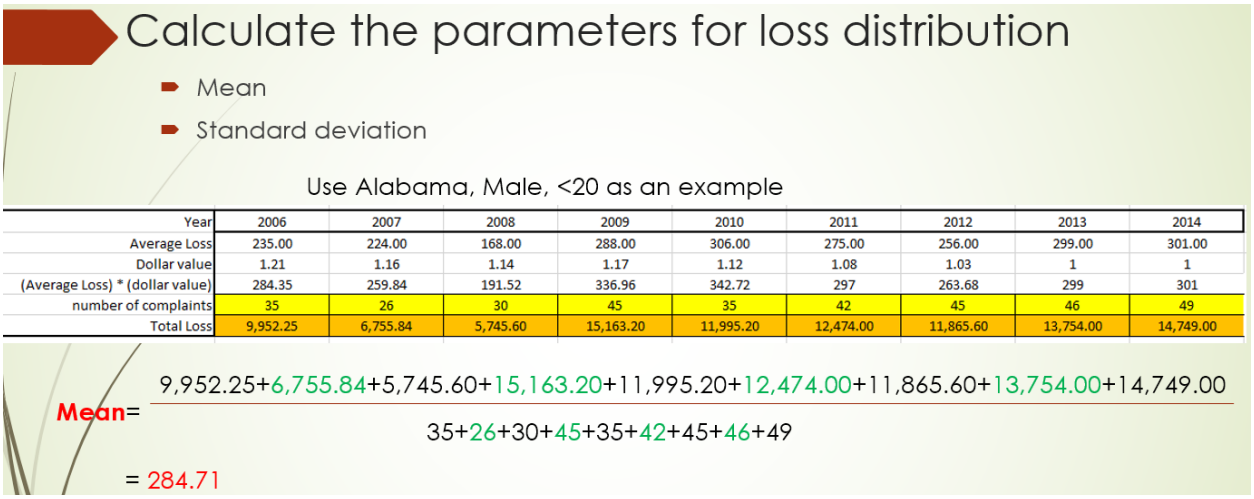


Figure 33

I calculated the mean of this group in 2014 by calculating the expected mean from 2006 to 2012.

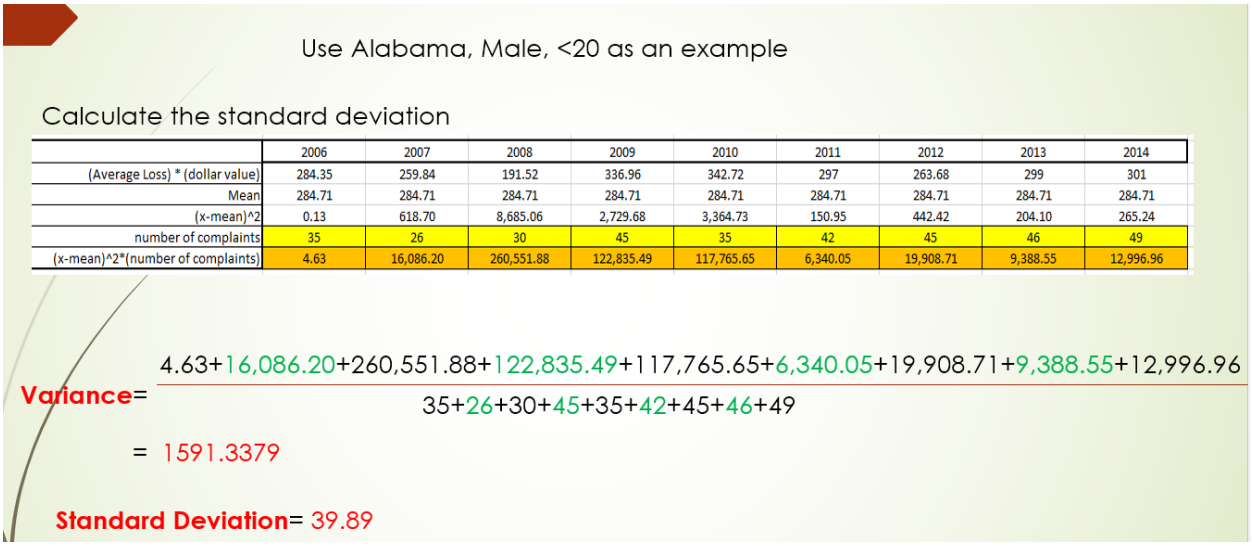


Figure 34

I calculated the standard deviation of this group in 2014 by calculating the expected standard deviation from 2006 to 2012

Then I used the same strategy to calculate parameters (mean and standard deviation) of the other 611 groups.

So far, I had built the loss models for 612 groups in 2014, and then I needed to use simulations to determine the cost of the insurance.

Use Matlab to simulation the loss model and determine the cost

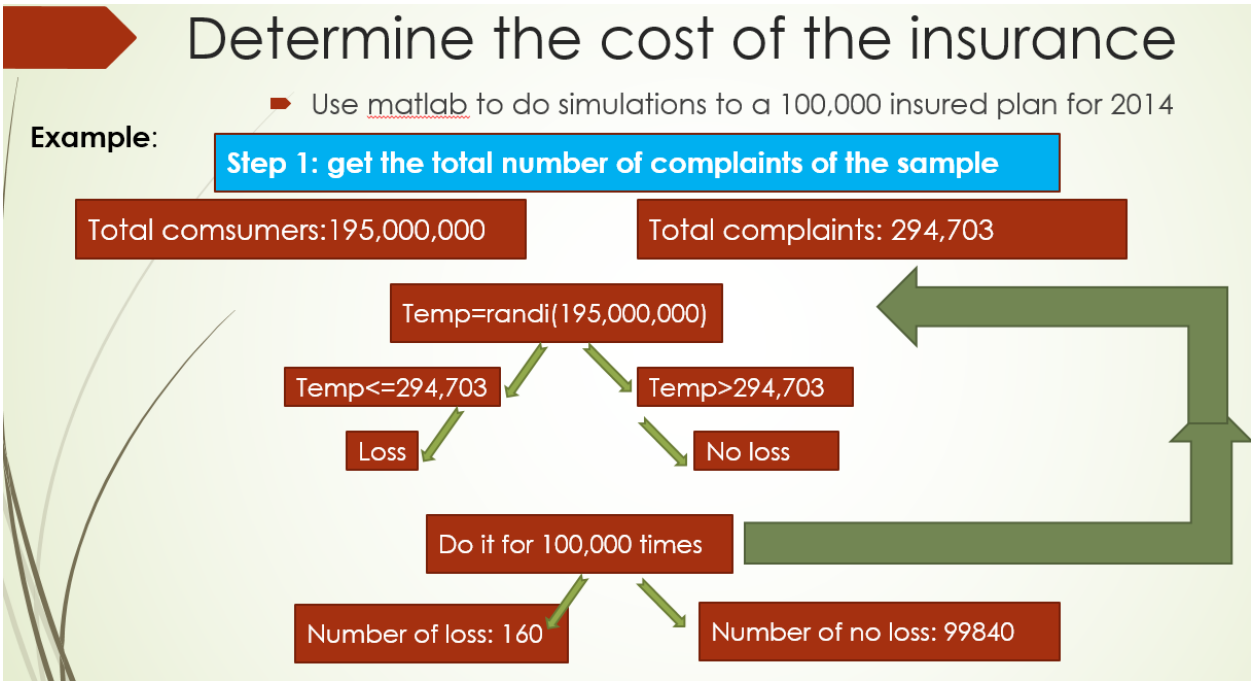


Figure 35

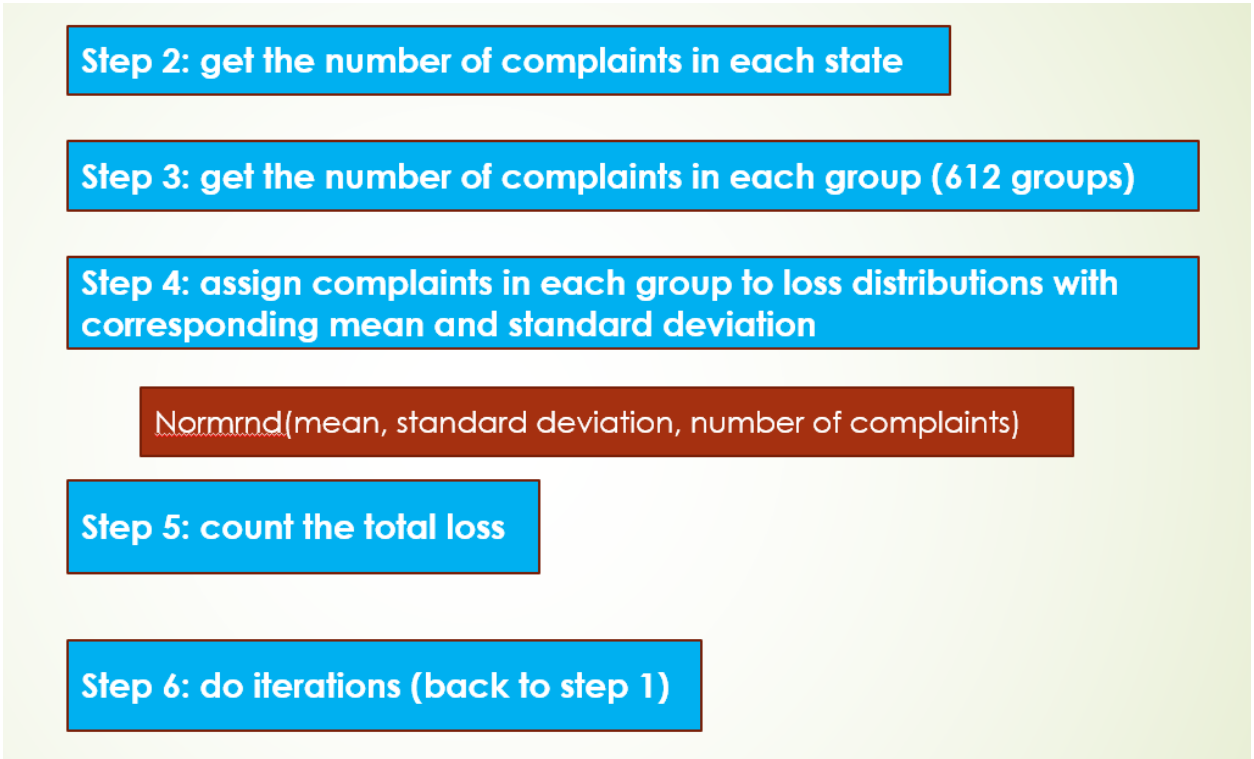


Figure 36

For each policyholder in a 100,000 insured plan, use the loss model (probability) to determine if the policyholder has a loss this year. If he/she does not have a loss, his/her loss will be 0 this year (no compensation to him/her). If he/she has a loss, I use the loss model to locate him/her into a state, a gender, an age group, and give him/her a loss using corresponding mean and standard deviation. For each iteration, there will be 100,000 losses (some have values and others are 0). I calculated the total of these 100,000 losses and the total number of policyholders who have a loss greater than 0. Then I did 10,000 iterations, and I calculated the average

100,000 insured each time for 10,000 times

	Average	Minimum	Maximum
number of loss (greater than 0)	151	102	198
loss ($)	272,102.62	171,335.64	401,444.97
cost per person ($)	2.72	1.71	4.01

No Max

price per person	min profit	max profit	average profit	profitable percentage
1.00	-301,444.97	-71,355.64	-172,102.62	0.00%
1.50	-251,444.97	-21,355.64	-122,102.62	0.00%
2.00	-201,444.97	28,644.36	-72,102.62	0.21%
2.50	-151,444.97	78,644.36	-22,102.62	20.21%
3.00	-101,444.97	128,644.36	27,897.38	85.42%
3.50	-51,444.97	178,644.36	77,897.38	99.79%
4.00	-1,444.97	228,644.36	127,897.38	99.99%
4.50	48,555.03	278,644.36	177,897.38	100.00%
5.00	98,555.03	328,644.36	227,897.38	100.00%
5.50	148,555.03	378,644.36	277,897.38	100.00%
6.00	198,555.03	428,644.36	327,897.38	100.00%
6.50	248,555.03	478,644.36	377,897.38	100.00%
7.00	298,555.03	528,644.36	427,897.38	100.00%
7.50	348,555.03	578,644.36	477,897.38	100.00%
8.00	398,555.03	628,644.36	527,897.38	100.00%
8.50	448,555.03	678,644.36	577,897.38	100.00%
9.00	498,555.03	728,644.36	627,897.38	100.00%

Figure 37

The data above shows the profit (average, minimum, and maximum) with different prices (regardless of expenses). For example, if the price is $3.50 per policyholder, the minimum profit will be $-51,444.97 (from the simulation), the maximum profit will be $178,644.36 (from the simulation), the average profit will be $77,897 (from the simulation), and 99.79% of the 10,000 iterations are profitable (3.5*100,000-total loss>0). As the data shows, $4.50 is the lowest safe price (from the simulation). Thus, if the insurance company sets up the price at more than $4.50 per person, it is very likely to earn money from this insurance product.

$4.50 would be a safe price from the simulation, but there is still the possibility for a policy holder to have an extremely large loss that could cause big trouble for the insurance company. Thus, the insurance company needs restrictions to reduce the risk. As my previous calculation shows, the maximum payment would be the best restriction to have. Then, I calculated the profit for different maximum prices.

Profit table with different maximum payment

No Max					Maximum Payment = 12,000				
price per person	min profit	max profit	average profit	profitable percentage	price per person	min profit	max profit	average profit	profitable percentage
1.00	-301,444.97	-71,355.64	-172,102.62	0.00%	1.00	-301,444.97	-71,355.64	-171,759.38	0.00%
1.50	-251,444.97	-21,355.64	-122,102.62	0.00%	1.50	-251,444.97	-21,355.64	-121,759.38	0.00%
2.00	-201,444.97	28,644.36	-72,102.62	0.21%	2.00	-201,444.97	28,644.36	-71,759.38	0.21%
2.50	-151,444.97	78,644.36	-22,102.62	20.21%	2.50	-151,444.97	78,644.36	-21,759.38	20.42%
3.00	-101,444.97	128,644.36	27,897.38	85.42%	3.00	-101,444.97	128,644.36	28,240.62	85.91%
3.50	-51,444.97	178,644.36	77,897.38	99.79%	3.50	-51,444.97	178,644.36	78,240.62	99.87%
4.00	-1,444.97	228,644.36	127,897.38	99.99%	4.00	-1,444.97	228,644.36	128,240.62	99.99%
4.50	48,555.03	278,644.36	177,897.38	100.00%	4.50	48,555.03	278,644.36	178,240.62	100.00%
5.00	98,555.03	328,644.36	227,897.38	100.00%	5.00	98,555.03	328,644.36	228,240.62	100.00%
5.50	148,555.03	378,644.36	277,897.38	100.00%	5.50	148,555.03	378,644.36	278,240.62	100.00%
6.00	198,555.03	428,644.36	327,897.38	100.00%	6.00	198,555.03	428,644.36	328,240.62	100.00%
6.50	248,555.03	478,644.36	377,897.38	100.00%	6.50	248,555.03	478,644.36	378,240.62	100.00%
7.00	298,555.03	528,644.36	427,897.38	100.00%	7.00	298,555.03	528,644.36	428,240.62	100.00%

	Maximum Payment = 3,000					Maximum Payment = 2,000				
profitable percentage	price per person	min profit	max profit	average profit	profitable percentage	price per person	min profit	max profit	average profit	profitable percentage
0.00%	1.00	-262,098.04	-61,992.23	-154,294.26	0.00%	1.00	-206,248.18	-43,033.10	-119,245.32	0.00%
0.00%	1.50	-212,098.04	-11,992.23	-104,294.26	0.00%	1.50	-156,248.18	6,966.90	-69,245.32	0.02%
0.35%	2.00	-162,098.04	38,007.77	-54,294.26	0.77%	2.00	-106,248.18	56,966.90	-19,245.32	15.77%
28.27%	2.50	-112,098.04	88,007.77	-4,294.26	43.22%	2.50	-56,248.18	106,966.90	30,754.68	94.57%
92.21%	3.00	-62,098.04	138,007.77	45,705.74	97.33%	3.00	-6,248.18	156,966.90	80,754.68	99.99%
99.95%	3.50	-12,098.04	188,007.77	95,705.74	99.99%	3.50	43,751.82	206,966.90	130,754.68	100.00%
100.00%	4.00	37,901.96	238,007.77	145,705.74	100.00%	4.00	93,751.82	256,966.90	180,754.68	100.00%
100.00%	4.50	87,901.96	288,007.77	195,705.74	100.00%	4.50	143,751.82	306,966.90	230,754.68	100.00%
100.00%	5.00	137,901.96	338,007.77	245,705.74	100.00%	5.00	193,751.82	356,966.90	280,754.68	100.00%
100.00%	5.50	187,901.96	388,007.77	295,705.74	100.00%	5.50	243,751.82	406,966.90	330,754.68	100.00%
100.00%	6.00	237,901.96	438,007.77	345,705.74	100.00%	6.00	293,751.82	456,966.90	380,754.68	100.00%
100.00%	6.50	287,901.96	488,007.77	395,705.74	100.00%	6.50	343,751.82	506,966.90	430,754.68	100.00%
100.00%	7.00	337,901.96	538,007.77	445,705.74	100.00%	7.00	393,751.82	556,966.90	480,754.68	100.00%

Figure 38

The average rate for the insurance premium to the compensation is about 0.23% in America. Therefore, letting the maximum price be 450 times the price per person would be a reasonable insurance system. For example, if the company sets up the price to be $5, the maximum payment could be $2250. This project focuses on figuring out the method of calculating the cost of the insurance.

Work with an anti-virus software company

Feasible analysis

From above, we can set up the approximate cost of insurance by simulating the situation in matlab. However, the risk set was not always stable, and the future risk may vary since this is a new kind of insurance and we do not have a large source of experienced data to predict a stable future risk. In addition, the market is also hard to predict and control since no companies have sold this kind of insurance. In order to solve this problem, I am thinking of aligning the online shopping insurance with a popular anti-virus software company. First, the software company's customers are supposed to be a certain group of people (the software company should have enough historical data to predict the future customers), and we could build a model from its shopping history. Second, we need the software company to obtain the policyholders' permission and detect the customers' online actions; when the customer is shopping with dangerous sellers, the software should suggest that the customer to stop the transaction. This approach could lower the customer's loss and also reduce the risk of our insurance to payout compensations. Third, the software should get the policyholders' permission and create a digital log about the customers' shopping history. This log could later provide the evidence to help regain the loss from the seller. The log can also serve as the evidence for our insurance to decide if the customer should be compensated and how much compensation the customer is going to receive.

Competitive Advantages Analysis

Corroborating with an anti-virus software company is a way to bring the traditional insurance company to e-commerce, which could help to extend the market for insurance companies. This cooperation will help the insurance company to reduce its risk, and the price could be more competitive or reasonable. Furthermore, the insurance company can analyze the market easily by working with the software company. The calculation (from the previous

method) for the cost of insurance will be more accurate by locating the policyholder among specific groups (the software company can provide the data of the customers' characteristics), and, in consequence, the future risk could be more stable than before. It will also be a good advertisement for the insurance company to enter the market of anti-virus software company. The success of this insurance product (Safeshop) could attract the customers to consider the insurance company's other products; after all, they will already have the anti-virus software on their computers or cell phones and it will be easy to let the customers view other product of this insurance company. In the future, the insurance company can personalize its users and introduce insurance products to corresponding type of users according to their online shopping activities.

Working together on the online shopping insurance will benefit the insurance company as well as the anti-virus software company. The previous prediction shows that the insurance product will have a high possibility of profitability. The anti-virus software technology can reduce more risk of the product and the software company will also be able to generate a good payback. The software company will not lose anything, even if this insurance product is not profitable, since the insurance company will be in charge of all of the compensations. Moreover, the anti-virus software company will be more competitive in its industry, as it guarantees its customers' loss that relates to online shopping fraud.

Service Agreement

In order to regulate the payment process and help to protect the legitimate interest of all parties (i.e., the insurance company, the anti-virus software company, and the policyholders), the service agreement will state, in detail, the payment conditions, non-payment conditions, and the payment process.

1. Compensation Standards:

1.1 When an insured customer encounters a loss under the protection of Safeshop (a insurance product combined with an anti-virus software) and makes a payment on internet browsers listed below:

…

…

…

This insurance will guarantee the insured customer a loss compensation with a maximum payment as X (use the previous simulation method, loss model and the user data from the anti-virus software company to calculate X).

1.2 Every insured customer can apply for a maximum of 3 losses per day, and the maximum annual compensation is X.

2. Payment Conditions

2.1 Open the protection prior of the payment

2.1.1 The insured must be a registered user of Safeshop and have Safeshop installed on his or her computer.

2.1.2 During the online shopping process, this insurance program (Safeshop) should be open. Further, the insured customer should be aware that this program will make notes to a product log, including the information below:

a. The starting time and ending time of this insurance program.

b. The websites that the insured customer visited during this shopping experience.

c. All of the documents downloaded by the insured customer during this shopping experience.

d. The condition of the insured's computer during this shopping experience.

e. The prompt message offered by the insurance program and the insured customer's reaction to each message.

2.1.3 The insured customer is using browsers listed on 1.1

2.2 Remedy after the loss

2.2.1 The insured customer should apply the loss and provide evidence to the appropriate police department or request for application assistance from the insurance program.

2.2.2 After the loss, the insured customer should apply for compensation through the process provided by the insurance program.

2.2.3 The insured customer should apply for compensation within 30 days after the loss.

2.2.4 The insured customer should keep the evidence. Do not reinstall the computer system, clean up the computer data, delete the product log, or delete the chatting record with the seller.

2.2.5 The insured customer should provide the required information to the insurance program, and the information should be real (if it is not, the insurance will not pay the loss).

2.2.6 The insured customer should agree to and authorize the use of the information provided to the insurance program, and the insurance program will use this information under the law of privacy.

2.2.7 The insured should provide the insurance program with the product log made by Safeshop, and he or she should understand that this log may contain privacy information. The insured customer should also authorize the insurance program to collect the evidence information from the product log.

3. Non-Payment Conditions

3.1 The insured customer makes the loss on purpose

3.1.1 The insured customer did not have prior protection as 2.1 described

3.1.2 The insured customer did not have remedy actions as 2.2 described

3.1.3 The insured customer provided fraudulent information to the insurance program in order to get the compensation.

3.1.4 The insured customer made the payment after noticing the warning messages provided by Safeshop.

3.1.5 The insured customer transfers money to others' personal accounts

3.1.6 The insured customer provides his or her bank account numbers or passwords to the seller.

3.1.7 After the insured customer had lost his credit or debit card, he or she did not notify the bank to block his or her card in time.

3.2 Special product

3.2.1 Transactions or recharges for online games.

3.2.2 Virtual goods

3.2.3 Medicine or health products

3.2.4 Obscene objects, video chat, pornography trading

3.2.5 Piracy or imitation

3.2.6 Smuggling or illegal goods

3.2.7 Network gambling or lottery

3.2.8 Other illegal products

3.3 Other conditions

3.3.1 Online part-time job

3.3.2 The insured customer was tricked by an online seller, but he or she made the payment offline

3.3.3 The loss is already covered by a third party

3.3.4 The required information and evidence is insufficient

4. Payment Process

4.1 After the loss, the insured customer files the application for compensation within 30 days

4.2 The insured customer then provides the accurate and complete information listed below:

a. basic information regarding his or her online shopping

b. accurate and valid personal information

c. related evidence and the product log

4.3 The insured customer submits the application

4.4 The insurance program will verify the information provided by the insured customer. The program will also respond to the insured as soon as possible (i.e., the program may require more information or evidence)

4.5 If the information and evidence provided by the insured is real and satisfies the payment conditions, then the insurance program will send the compensation to the insured within 30 days.

Conclusions

The goal of this project was to develop a new type of insurance—online shopping insurance—and the project focused on developing a method to calculate the cost of such a product according to historical data (ic3's annual reports from 2006 to 2010), and analyze the ways to reduce the risk.

The historical data formed linear trends after they were divided into 612 different groups for each year. Then, the data for 2014 was predicted by using the linear least square approach. Matlab simulations showed that maximum payment played a better role than the other two methods for reducing the risk.

When applying the loss model to a 100,000-policyholder insurance plan, the matlab simulation showed that the insurance product had an average cost (cost per policyholder) of $2.72, a minimum cost (cost per policyholder) of $1.71, and a maximum cost (cost per person) of $4.01. If the insurance company establishes the price higher than $4.50, then this insurance product would have a high possibility of generating a profit (regardless of expenses). The tendency for this insurance product (online shopping insurance) to be profitable would count as a bargaining chip for the insurance company to persuade an anti-virus software company to team up to make this project a reality.

In conclusion, Safeshop, an online shopping insurance product, could be a valuable investment for insurance vendors and anti-virus software companies. It can help online shoppers to transfer risk to the insurance company, reduce the government's financial burden, aid the traditional insurance companies in entering the e-commerce market, and build a safe online shopping environment.

Works Cited

"." FBI. FBI, 8 May 2013. Web. 30 Apr. 2014. <http://www.fbi.gov/>.

"Monthly & Annual Retail Trade." , Main Page. N.p., n.d. Web. 30 Apr. 2014. <http://www.census.gov/retail/>.

"2006 Internet Crime Report." . ic3.gov, 1 Jan. 2006. Web. 30 Apr. 2014. <http://www.ic3.gov/media/annualreport/2006_IC3Report.pdf>.

"2007 Internet Crime Report." . ic3.gov, 1 Jan. 2007. Web. 30 Apr. 2014. <http://www.ic3.gov/media/annualreport/2007_IC3Report.pdf>.

"2008 Internet Crime Report." . ic3.gov, 1 Jan. 2008. Web. 30 Apr. 2014. <http://www.ic3.gov/media/annualreport/2008_IC3Report.pdf>.

"2009 Internet Crime Report." . ic3.gov, 1 Jan. 2009. Web. 30 Apr. 2014. <http://www.ic3.gov/media/annualreport/2009_IC3Report.pdf>.

"2010 Internet Crime Report." . ic3.gov, 1 Jan. 2010. Web. 30 Apr. 2014. <http://www.ic3.gov/media/annualreport/2010_IC3Report.pdf>.

"2011 Internet Crime Report." . ic3.gov, 1 Jan. 2011. Web. 30 Apr. 2014. <http://www.ic3.gov/media/annualreport/2011_IC3Report.pdf>.

"2012 Internet Crime Report." . ic3.gov, 1 Jan. 2012. Web. 30 Apr. 2014. <http://www.ic3.gov/media/annualreport/2012_IC3Report.pdf>.