Modeling Motorcycle Insurance Rate Reduction due to Mandatory Safety Courses

Shujuan Huang[1], Vadiveloo Jeyaraj[1], Valdez Emiliano[1], Garry D. Lapidus[2]

[1]Mathematic Department, University of Connecticut, Storrs, CT 06269, USA

[2] Injury Prevention Center, Connecticut Children's Medical Center, University of Connecticut School of Medicine, Hartford, CT 06106, USA

Abstract

Alarming statistic has indicated that the risk of fatality associated with motorcycle crashes far exceeds that of automobiles; hereby Connecticut introduced mandatory motorcycle safety training. This paper develops a unifying framework to quantify the effectiveness of such mandatory programs and to translate this in terms of a possible insurance rate reduction. Overall Discount Rate Estimation and Individual Discount Rate Adjustment are achieved by nonlinear optimization and an Integer-Valued Autoregressive (INAR) model, respectively. A heterogeneity factor is injected into the model to assess the impact of the training programs. Finally, numerical Illustrations are given with data drawn from Connecticut.

1 Introduction

1.1 Research Purpose

1.1.1 Intense call on Insurance Industry Involvement from Public Transportation and Health Department

In 2008, 5,290 U.S. motorcyclists were killed and 96,000 injured[12]. Since 1998, there has been a significant increase in deaths and non-fatal injuries. Per vehicle mile traveled, motorcyclists are about 37 times more likely than passenger car occupants to die in a motor vehicle crash and 9 times more likely to be injured[12]. Motorcycle-related deaths account for 14 percent of total traffic fatalities in 2008, although motorcycles only made up about 3 percent of registered vehicles[12]. More and more studies have indicated that the risk of driver and passenger fatality associated with motorcycle crashes far exceeds that of automobile crashes. This alarming statistic has caused increased attention from state motor vehicle departments, state health departments and the insurance industry to pursue efforts to introduce programs to dramatically reduce motorcycle accidents.

In the "National Agenda for Motorcycle Safety" which is supported by Motorcycle Safety Foundation (MSF) and National Highway Traffic Safety Administration (NHTSA) of US Department of transportation (DOT), it is strongly recommended that the insurance industry should collect, organize, analyze, and distribute motorcycle-specific loss data to better understand safety issues, and develop guidelines to tie approved training, licensing, and safe-riding practices to premium reductions. However, at present practice, insurers employ limited avenues to enhance and encourage motorcycle safety.

Motorcycle insurers are not currently required to provide motorcycle-specific loss data for analysis or use a safety-related database to guide insurance policy. It also should be noted that most research or projects which focus on evaluating the effectiveness of motorcycle safety course systematically and comprehensively were completed before 2000[2,6,7,8,10,11]. These studies have shown mixed results on effectiveness, due to different kinds of methodological issues. In addition, there is no training indicator variable included in current national traffic accident databases. While some insurers have both related claim records and information about the policyholder, and whether they have taken the training or not, it would be extremely helpful to evaluate these the efficacy of motorcycle rider training programs if we could employ more resources in the insurance industry.

1.1.2 Reduce Insurers' Own Losses by Supporting Certain Responsible Riding Practices with Incentives.

Currently, in some states, most motorcycle insurance companies offer up to a 10% discount within the last three years, for the successful completion of the Basic or Experienced Motorcycle Safety Course. The courses are uniformly designed and guided by MSF, and the content should be nearly consistent. However, the discount rate varies greatly from company to company, and from state to state. From the perspective of the insurers' own interest, developing a quantitative methodology to evaluate the effectiveness of the Motorcycle Safety Courses and the insurance discount rate， would enhance motorcycle safety and optimize insurers' risk management as well.

Table.1 Facts about current motorcycle insurance discount rate for safety course

Insurance Company	Discount rate	Details[1]
PROGRESSIVE	NA[2]	Safety Course – Completing an approved safety course could earn you a discount.
GEICO	10%	10% discount for completing a Motorcycle Safety Foundation or Military Safety Course
Allstate	5%	Save 5% if you've voluntarily passed a Motorcycle Safe Driving in the past 36 months.
USAA	5%	approved safety course within the last three years
FOREMOST	NA	Motorcycle safety course discount
Nationwide	up to 5%	Save up to 5 percent on your motorcycle insurance when you complete an approved safety course.
MARKEL	NA	Safety Course Discount
Dairyland Cycle	NA	Motorcycle safety course completion
Rider	No discount	

[1] From official website of insurance companies

[2] NA means the specific value for discount rate is not disclosed directly on the website. The customers need to consult the agent case by case.

1.2 Facts about Motorcycle Safety Course

Nearly all the approved safety courses recognized by different insurance companies rely on the Motorcycle Safety Foundation (MSF) RiderCourses, which are adopted by most states DMV. The Motorcycle Safety Foundation is an internationally recognized developer of the comprehensive, research-based, Rider Education and Training System (MSF RETS), which is a not-for-profit organization sponsored by BMW, BRP, Ducati etc.

There are different levels of Rider Courses, for example: Basic RiderCourse(BRC), Basic RiderCourse 2(License waiver, skill practice), Street RiderCourse 1, Returning Rider Basic Rider Course, 3-Wheel Basic RiderCourse (3WBRC), Scooter Basic RiderCourse (SBRC), Street RiderCourse 2 (SRC2), Experienced RiderCourse (ERC). Some of courses are just being introduced or in development or design phase. Different states and different driving schools approved by MSF may offer levels of instruction or give the courses different names. For example, in Waterbury, CT, the courses offered by the Rider Education Program include the Basic RiderCourse, Intermediate RiderCourse, and the Experienced RiderCourse. In New York State, they designate the intermediate Rider Course as the Basic RiderCourse 2. However, the content of the Basic RiderCourse 1 and 2 (sometimes called Intermediate RiderCourse) are similar. Both courses are for those motorcycle riders who do not have a license yet. While the advanced RiderCourse are for those who have been riding for some time. In this paper we will only consider two categories: BRC (Basic RiderCourse) and ARC (Advanced RiderCourse).

1.3 Introduction to Automobile Insurance and Priori Rating System

In an insurance portfolio, the potential risks exposed by policyholders vary; specifically for automobile insurance, the likelihood of having crashes varies among the insured drivers. One of the main tasks of actuaries is to fairly allocate the burden of baring the potential losses among policyholders, which is materialized by quantitative analysis to specify individual risks and thus to determine the premiums. This procedure is called pricing or rate-making. There are two main phases involved. A base premium is determined when the policy is issued, and then the premium will be adjusted by discounts or surcharges as the policy is carried out. A discount for the motorcyclists who have taken the safety course designed by MSF would qualify for a discount. In theory and practice, motorcycle riders should pay a premium corresponding to his/her own risk, and evaluate this discount rate according to the effectiveness on risk reduction of motorcycle rider safety training program.

1.4 Why Research Studies about the Effectiveness have Shown Mixed Results

The results of research studies looking at the effectiveness of rider training have shown mixed results[2,6,7,8,10,11]. Most of the studies reviewed a training program that essentially consisted of a single course. Most state governments and insurance company involvement in the U.S. are through the licensing function, and therefore, limited primarily to a basic novice course. In addition, MSF Rider Education and Training System

(RETS) is expanding in breadth and depth to meet the growing needs of current and prospective riders all the while. For example, the *Street RiderCourse* was just introduced in the last year (2010), and there may be a time lag to show program effectiveness. Some courses are still under the process of developing and have not been introduced yet. These may lead to a fallacy of a single training course serving as an in-total countermeasure.

2 Effectiveness of Different Levels of Motorcycle Safety Courses

2.1 Definition of effectiveness

Effectiveness[3] is a measure of the extent to which a specific intervention, procedure, regimen, or service, when deployed in the field in routine circumstances, does what it is intended to do for a specific population. In other words, it means doing "right" things, i.e. setting right targets to achieve an overall goal (the effect). Then we need to define our overall goal first. To put it simply, if the goal is to assure the minimum riding skills for initial entry into the motorcycling environment, then we can say MSF safety course has achieved at a 85-90% success rate in basic courses according to the records of training schools. However, in most cases, we need to consider a more comprehensive goal of safety courses which is to determine if motorcycle rider safety training courses have any impact on reducing the frequency of motorcycle crashes, injuries, and insurance claims. It may also include: quality education and training, knowledge, skills, attitude, habits, values, risk management skills, self-awareness and self-assessment.

Since most motorcycle insurance companies offer up to a 10% discount for the successful completion of the Basic or Experienced Motorcycle Safety Course, the effectiveness of the safety course would directly affect the costs of insurance companies. From the point of view of Motorcycle insurance pricing, we could define actuarial-effectiveness as follows:

Definition:

Actuarial-Effectiveness for Motorcycle Safety Training**:** The average reduction ratio of incurred loss per unit exposure (or average claim frequency) to insurance company that claimed by a population of untrained motorcyclists if all would have taken the safety training, without any change else (e.g., weather, motorcycle physical condition, transportation environment). This is essentially the same as the expected risk reduction for a motorcyclist to attend the safety training from non-attending state. For example: Actuarial-Effectiveness of 10 percent means that an insurance company can reduce their incurred loss for one policy holder (or frequency of claims) by 10% simply by convincing that policy holder to attend the safety training course.

[3] Last JM. A Dictionary of Epidemiology (4th Edition). New York: Oxford University Press 2000.

2.2 Formulation of Actuarial-Effectiveness

Note that only those Advanced Motorcycle Safety Course learners have previous records, and the novice training course learners don't, which means that they have no previous claim record. Therefore, there should be different formula for different courses' effectiveness. I will begin by addressing the situation where the data is free from censoring or truncation.

2.2.1 Effectiveness for Basic Motorcycle Safety Course

For the Basic Motorcycle Safety Course, we need to compare the reduction on incurred loss per unit of exposure after training with those without training since before training the BRC learners don't have a license yet. In particular, the estimated effectiveness makes sense only when the chosen samples have similar characteristics such as age, gender, motorcycle models, years riding, miles ridden per year and primary purpose of riding (commuting, recreation, etc.). Hence matched-pair approach could be exploited here to calculate the effectiveness for basic motorcycle safety course. Usually, for insurers, we have the number of incurred claims, unites of exposures, dollars of incurred losses per year, as well as whether they have taken the training or not (if yes, the kind of course they have taken). The corresponding claim data about both claim frequency and severity of the policyholder should be also available. For the matched-pair samples, suppose we have obtained the following summary data about exposures, claim count and incurred losses.

Here the timeline should be based on the motorcyclists' training year. We denote the year they took training as $t=0$, one year after they took the training as $t=1$, and two years after they took training as $t=2$. While for those who have not taken any training, we don't need to consider such time restriction.

Table 2 Matched-pair sample summary for BRC

Types of policyholders when loss occurs	Exposures for year t	Total Claim Count	Total Incurred Loss
Have taken BRC[4] before	e_{Bt}	C_{Bt}	L_{Bt}
No training	e_N	C_N	L_N

If we define the effectiveness by the measure of reduction in incurred loss, the effectiveness[5] for year t of BRC would be: $\rho_B = \left(1 - \dfrac{\sum_{t=0}^{2} L_{Bt} \Big/ \sum_{t=0}^{2} e_{Bt}}{L_N / e_N}\right)_+$

[4] For those have taken both BRC and ERC in history, only consider the most recent one, in other words, the ERC.

[5] When the value of ρ_{Bt} is negative, we take it as 0, mean it is not effective at all

If we define the effectiveness by the measure of reduction in claim frequency, the effectiveness for year t of BRC would be:

$$\rho_B = \left(1 - \frac{\sum_{t=0}^{2} C_{Bt} \Big/ \sum_{t=0}^{2} e_{Bt}}{C_N / e_N} \right)_+$$

2.2.2 Effectiveness for Advanced Motorcycle Safety Course

For the Advanced Motorcycle Safety Course, we can compare the claim data for the same policyholder before and after they took the course. The timeline should also be based on the motorcyclists' training year. For different policyholders, however, their training years may be different. Here let n_i denote the number of years for policyholder i's records before training, and m_i denote the number of years after.

Table3 Claim history statistics for motorcyclists who took the ERC

For those who has taken ERC	Exposures	Total Claim Count	Total Incurred Loss
Before taken ERC	e_E	$C_{E,i1}$	$L_{E,i1}$
After taken ERC	e_E	$C_{E,i2}$	$L_{E,i2}$

If we define the effectiveness by the measure of reduction in loss, the effectiveness of ERC would be:

$$\rho_E = \frac{1}{e_E} \sum_{i=1}^{e_E} \left(1 - \frac{L_{E,i2} / m_i}{L_{E,i1} / n_i} \right)_+$$

If we define the effectiveness by the measure of reduction in claim frequency, the effectiveness for year t of BRC would be:

$$\rho_E = \frac{1}{e_E} \sum_{i=1}^{e_E} \left(1 - \frac{C_{E,i2} / m_i}{C_{E,i1} / n_i} \right)_+$$

3 Motorcycle Insurance Rate Reduction Modeling

In the following sections, we will use the effectiveness by the measure of reduction in claim frequency, that is the (frequency part of the) pure premium[6]. Firstly, we will estimate the overall discount rate using insurance claim data drawn from states where such mandatory programs have been introduced. Secondly, similar to the bonus-malus scheme for each policy holder, when we need to determine the specific discount rate for each person, we would consider both the policy holder's claim history and overall discount rate.

[6] The complete pure premium includes also the cost of the claim. It is equal to the frequency part times the expected cost per claim, when cost per claim and claim occurrence are independent.

3.1 Reflection on Current Training Discount Rate Policy

As shown in table 1 of section 1.1, some insurance companies use unified discount rate, such as 5% in Allstate and 10% in GEICO, while other insurance companies use flexible policies for different policy holders, such as up to 5% in Nationwide or negotiate case by case in PROGRESSIVE. It should also be noted that some insurers require the safety course should be taken within 3 years like Allstate and USAA, while others not.

Based on those differences, from the point of view of risk management and profit maximization of insurers, we will first evaluate whether we need to add the constraints that the safety course should be taken within 3 years, then determine what specific value or upper bound we should use.

3.2 Evaluation about the "3-year" Constraints

This could be easily carried out by check the effectiveness of the BRC and ERC.

For BRC, as we discussed before $\rho_{Bt} = \left(\frac{C_N / e_N - C_{Bt} / e_{Bt}}{C_N / e_N} \right)_+$, where we denote the year they took training as $t = 0$, one year after they took the training as $t = 1$, two years after they took training as $t = 2$, and so on. Then we can compare ρ_{B4} with ρ_{B1}, ρ_{B2} and ρ_{B3} by a simple statistical test, if there are significant difference between them, we could add the *3-year* constraint, otherwise, it is unnecessary. Meanwhile, we could evaluate other possible *k-year* policy using similar methods.

For the ERC, as we discussed before $\rho_E = \frac{1}{e_E} \sum_{i=1}^{e_E} \left(1 - \frac{C_{E,i2} / m_i}{C_{E,i1} / n_i} \right)_+$. In order to evaluate the *3-year* policy, we could fix n_i as 1, 2, 3, 4 or more. Then compare whether there is any significant statistical difference when n_i is larger than 3. If it is, we suggest keep the *3-year* policy, otherwise not.

3.3 Overall Discount Rate Estimation Using Past Insurance Claim Data

Firstly, we could use past insurance claim data to estimate the overall discount rate, considering both the training effect on pure premium and the demand. Then based on this estimated overall discount rate, discount for individual motorcyclist could be further adjusted according to their own claim history.

3.3.1 Incentive Adjustment on Discount Rate

Since from January 1st, 2011, Connecticut requires motorcyclists of all ages who want to get the license to take the BRC, insurance companies don't need to use discount rate to attract the customers but should offer basic discount based on the reduced risk by training. While the ERC is voluntary, then it is possible that someone took the course just for the discount, and the riders' view on effectiveness of the ERC also counts.

Therefore, the number of policyholder who has taken ERC (within three year) should be a function of discount rate r, the effectiveness of the course ρ_E and the total demand N_t. This paper denotes that function as $e_{Et} = f\left(r, \rho_E, N_t\right)$. For most states where the BRC is not required, an insurer could also consider the incentive adjustment on BRC discount rate to increase the number of policyholder to reduce the total risk and increase profit. There are three factors which can affect total profit and risk management of insurance companies as follows:

Table 4: Three factors that affect profits of insurer

Type	variable	Impact on insurance company profit
Reduction in individual premium because of the discount offered to customers	Discount rate r	-
Possible Reduction in Claim Cost because the training effect in Risk Reduction	Effectiveness ρ_E and ρ_B	+
Possible increasing number of policy holders because of the discount incentives or state government requirement on training	Exposures e_{Et}, e_{Bt}, e_{Nt}	+/-

3.3.2 Overall Discount Rate for ERC and BRC learners

As we discussed before, for the ERC learners, we can compare their claim history before and after the training. If we have derived (the frequency part of) pure premium before training PP_1 and after training PP_2 separately, we could simply calculate the discount rate as $r = \left(PP_1 - PP_2\right)/PP_1$. Meanwhile, comprehensively considering the pure premium and demand, we will also consider introducing the adjustment factor κ. Suppose among all the policyholders who have taken the safety course for year t, the proportion for BRC is ∂ and the proportion for ERC is $1-\partial$. Once we have estimated r_E and r_B based on the changes on pure premium, the unified discount rate r could be $r = \kappa\left(\partial r_E + \left(1-\partial\right) r_B\right)$. We solve the optimization problem with following steps.

Step 1: Estimate r_A and r_B based on the changes on pure premium
Actually, here we could directly use the value of ρ_E and ρ_B to initially estimate r_E and r_B because the formula for ρ_E and ρ_B exactly reflects the change on (severity/ frequency part of) pure premium using the insurance claim data.

Step 2: Estimate the total in-force exposure at a certain time t
For current year's exposure, we could use the basic methods in ratemaking to estimate the total in-force exposures at a certain time, let's denote it as N_t. Then we could have $e_{Et} + e_{Bt} + e_{Nt} = N_t$

For future year's total exposure, we could use time series forecasting methods to estimate them. It should be noted that we need to consider the government mandatory policy's affect on the number of motorcyclists who have taken the course in the course of analysis.

Step 3: Formulate e_{Et} and e_{Bt}

e_{Bt} should depend on the policies of different States DMV. Some states require all-age motorcyclists to take the BRC if they want to get the license (e.g., Connecticut, Texas); while others require the motorcyclists under 18 (or 16, 21) to take the course. In this paper, we will study the former situation. Since we just introduced the policy in 2011, we could estimate e_{Bt} by the entire new license issuing pattern in the past years by the method of time series analysis.

As we discussed before, the number of policyholder who have taken ERC (within three years) should be a function of discount rate r, the effectiveness of the course ρ_E and the total demand N_t, that is $e_{Et} = f\left(r, \rho_E, N_t\right)$. Objectively, the function f should be monotone non-decreasing function of both r and ρ_E when N_t is given. We assume ρ_E to be a constant in a certain period. ρ_E could be estimated by the data in section 2.2. Here we treat r (or adjustment factor κ) as a decision variable in our programming.

In fact, the function f could be treated as the utility function on the effectiveness of the training and the insurance discount for customers. Similar to the commonly used Exponential Utility[7] in insurance industry, we assume

$$e_{Et} = f\left(r, \rho_E, N_t\right) = \left(N_t - e_{Bt}\right)\left(1 - e^{-(\alpha r + \beta \rho_E)}\right) \tag{3.1}$$

where the factor α, β can be determined subjectively or objectively. To put it simply, we could let both α and β be 0.5.

Step 4: Final Programming

$$\text{Objective: } \underset{r}{Arg} \text{ Max } \rho_E e_{Et} \frac{L_{E,t-1}}{e_{E,t-1}} + \rho_B e_{Bt} \frac{L_{N,t-1}}{e_{N,t-1}} - r\left(e_{Et} + e_{Bt}\right) P_t \tag{3.2}$$

St. $0 \le r \le 15\%; k > 0$

$$r = \kappa\left(\partial r_E + \left(1 - \partial\right) r_B\right) \tag{3.3}$$

$$e_{Et} + e_{Bt} + e_{Nt} = N_t \tag{3.4}$$

$$e_{Et} = \left(N_t - e_{Bt}\right)\left(1 - e^{-(\alpha r + \beta \rho_E)}\right)$$

[7] "Utility-Theoretic Underwriting," http://www.casact.org/sections/care/0905/handouts/halliwell.pdf

Where N_t, e_{Bt} need to be estimated or calculated by past data. $L_{A,t-1}, e_{A,t-1}, L_{N,t-1}, e_{N,t-1}$ are based the previous years' data file. ρ_E, ρ_B are estimated in section 2.2, the same for r_E and r_B. This is a nonlinear programming problem. We could use MATLAB to solve it. We could get the optimized value for κ firstly, then the value for r easily followed.

3.4 Individual Discount Rate Adjustment by Personal Claim History

Since we have estimated the overall discount rate r in section 3.3, now we aim at determining the specific discount rate for each person, which is similar to the bonus-malus scheme for each policy holder.

3.4.1Discount Rate for ERC learners

In this paper, the Integer-Valued Autoregressive (INAR) method will be used to model individual annual claim count in consecutive policy years. Al-Osh and Alzaid (1987) proposed what they have called an integer-valued first order autoregressive (INAR(1)) model. Later, Gourieroux and Jasiak (2004) applied it to car insurance in bonus-malus system design. They compared the advantages of INAR than the traditional negative binomial approach. Zhang(2009) generalized the INAR to dynamic heterogeneity with applications in automobile insurance. Here we will use the similar INAR(1) model, but interpret the heterogeneity Θ_i as training impact factor.

Let's consider the policy holder i and denote $N_{i,t}$ the number of claims at year t submitted by this individual. We assume $N_{i,-1},...,N_{i,t},N_{i,t+1}$ are independent conditional on an unobservable heterogeneity factor Θ_i. We assume that the heterogeneity Θ_i is a time independent random variable and follows a Gamma distribution $\gamma(a,(1-r)/a)$. Here the parameters are design to ensure the expectation of Θ_i to be $1-r$; where r is the overall discount rate estimated in section 3.3.

$$N_{i,t} = B_t(p) \circ N_{i,t-1} + \varepsilon_{i,t} \tag{3.5}$$

$$\varepsilon_{i,t} \left| (\Theta_i = \theta_i) \right. \sim Poission(\theta_i \lambda_i) \tag{3.6}$$

$$\Theta_i \sim Gamma(a, \frac{1-r}{a})^{8} \tag{3.7}$$

$$E[\Theta_i] = 1-r \tag{3.8}$$

Where $\{\varepsilon_{i,t}\}_{t=-1}^{\infty}$ is a sequence of random variables taking nonnegative integer values; $B_t(p)$ is the so-called Binomial thinning factor, which is independent of the error term and defined by $B_t(p) \circ N_{i,t-1} = \sum_{j=0}^{N_{i,t-1}} U_j$, U_j is a sequence of i.i.d. Bernoulli random

[8] Recall: For gamma distribution $Gamma(k,\theta)$, where k is the shape parameter and θ is the scale parameter. Mean $\mu = k\theta$

variables with autoregressive parameter p .The mixture operation '$\circ$' is called binomial thinning.

Note:

(i) Intuitively, $N_{i,t}$ is introduced into two parts, one is the lagged claim counts from previous years $B_t\left(p\right)\circ N_{i,t-1}$, the second part is the newly arrived claim counts $\varepsilon_{i,t}$. Lagged claim counts in the first part are introduced because during the loss settlement period, which can be many years in duration, additional facts regarding individual claims and trends often will become known (including unpaid, and often unreported, losses to their ultimate settlement values)

(ii) For expository purpose we focus on the autoregressive process of order 1, but the approach is easy to extend to higher autoregressive orders.

(iii) $N_{i,t}$ is the average annual claim frequency for policy holder i during the year t (We denote the year before they took ERC as $t=-1$, the year they took training as $t=0$, one year after they took the training as $t=1$, two years after they took training as $t=2$, etc.) Therefore, here the index t is not indicating the exact year, but a relative time compare to the time when the motorcyclists took the training. Then $N_{i,t}$ is the corresponding statistics. Here we only consider one year before the training because there should be no training impact for consecutive years before the training, hence mismatching with formula (3.5) and (3.6).

At year t, the (frequency part of) pure premium is

$$\begin{aligned}
P_{t,i} &= E\left[N_{i,t+1}\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right] \\
&= E\left[E\left[N_{i,t+1}\middle|N_{i,-1},N_{i,0},...,N_{i,t},\Theta_i\right]\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right] \\
&= E\left[pN_{i,t}+\lambda_i\Theta_i\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right] \\
&= pN_{i,t}+\lambda_i E\left[\Theta_i\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right]
\end{aligned}$$

Where λ_i is predetermined by the basic ratemaking category of such policy holder.The risk on the count variable $N_{i,t+1}$ can be measured by

$$\begin{aligned}
R_{t,i} &= V\left[N_{i,t+1}\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right] \\
&= E\left[V\left[N_{i,t+1}\middle|N_{i,-1},N_{i,0},...,N_{i,t},\Theta_i\right]\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right] \\
&\quad +V\left[E\left[N_{i,t+1}\middle|N_{i,-1},N_{i,0},...,N_{i,t},\Theta_i\right]\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right] \\
&= E\left[p(1-p)N_{i,t}+\lambda_i\Theta_i\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right] \\
&\quad +V\left[pN_{i,t}+\lambda_i\Theta_i\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right] \\
&= p(1-p)N_{i,t}+\lambda_i E\left[\Theta_i\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right]+\lambda_i^2 V\left[\Theta_i\middle|N_{i,-1},N_{i,0},...,N_{i,t}\right]
\end{aligned}$$

As suggested by many investigations, the period that the safety courses play a key role to reduce risk is within 3 years after the course taken time (Of course it is easy to generalize 3 to any other number k if necessary), here we only need to consider the short claim history $t=-1,0,1,2$, that pertains to a customer with a seniority for up to 3 years or new customers with similar history.

Proposition 3.1 Conditional Distribution of Θ_i

(i) For $t=-1$, the conditional distribution of Θ_i is $\gamma(a,\frac{1-r}{a})$

(ii) For $t=0$, the conditional distribution of Θ_i given $N_{i,-1}$ is

$$\gamma\left[a+N_{i,-1},1/\left(\frac{a}{1-r}+\frac{\lambda_i}{1-p}\right)\right]$$

(iii) For $t=1$, the conditional distribution of Θ_i given $N_{i,-1}$, $N_{i,0}$ is

$$\frac{\sum_{z_2=0}^{\min(N_{i,-1},N_{i,0})}\pi\left(z_2,N_{i,-1},N_{i,0}\right)\gamma\left[a+N_{i,-1}+N_{i,0}-z_2,1/\left(\frac{a}{1-r}+\lambda_i+\frac{\lambda_i}{1-p}\right)\right]}{\sum_{z_2=0}^{\min(N_{i,-1},N_{i,0})}\pi\left(z_2,N_{i,-1},N_{i,0}\right)}$$

Where

$$\pi\left(z_2,N_{i,-1},N_{i,0}\right)=C_{N_{i,-1}}^{z_2}\left(\frac{p}{1-p}\right)^{z_2}\frac{\tau\left(a+N_{i,-1}+N_{i,0}-z_2\right)\left(\frac{a}{1-r}+\lambda_i+\frac{\lambda_i}{1-p}\right)^{a+N_{i,-1}+N_{i,0}-z_2}}{\lambda_i^{z_2}\left(N_{i,0}-z_2\right)!}$$

(iv) For $t=2$, the conditional distribution of Θ_i given $N_{i,-1}$, $N_{i,0}$, $N_{i,1}$ is

$$\frac{\sum_{z_3=0}^{\min(N_{i,1},N_{i,0})}\sum_{z_2=0}^{\min(N_{i,-1},N_{i,0})}\pi\left(z_2,z_3,N_{i,-1},N_{i,0},N_{i,1}\right)\gamma\left[a+N_{i,-1}+N_{i,0}+N_{i,1}-z_2-z_3,1/\left(\frac{a}{1-r}+2\lambda_i+\frac{\lambda_i}{1-p}\right)\right]}{\sum_{z_3=0}^{\min(N_{i,1},N_{i,0})}\sum_{z_2=0}^{\min(N_{i,-1},N_{i,0})}\pi\left(z_2,z_3,N_{i,-1},N_{i,0},N_{i,1}\right)}$$

Where $\pi\left(z_2,z_3,N_{i,-1},N_{i,0},N_{i,1}\right)=$

$$C_{N_{i,0}}^{z_3}C_{N_{i,-1}}^{z_2}\left(\frac{p}{1-p}\right)^{z_2+z_3}\frac{\tau\left(a+N_{i,-1}+N_{i,0}+N_{i,1}-z_2-z_3\right)\left(\frac{a}{1-r}+2\lambda_i+\frac{\lambda_i}{1-p}\right)^{a+N_{i,-1}+N_{i,0}+N_{i,1}-z_2-z_3}}{\lambda_i^{z_2+z_3}\left(N_{i,0}-z_2\right)!\left(N_{i,1}-z_3\right)!}$$

Proof:

Case $t=0$

The joint distribution of $N_{i,-1}$ and Θ_i is

$l(N_{i,-1},\Theta_i)=l\left(N_{i,-1}\middle|\Theta_i\right)l\left(\Theta_i\right)$

$$= \exp(-\frac{\lambda_i \Theta_i}{1-p}) \left[\frac{\lambda_i \Theta_i}{1-p}\right]^{N_{i,-1}} \frac{1}{N_{i,-1}!} \Theta_i^{a-1} \exp\left(-\frac{a\Theta_i}{1-r}\right) \frac{a^a}{\tau(a)(1-r)^a}$$

$$\sim \Theta_i^{N_{i,-1}+a-1} \exp\left(-\Theta_i \left(\frac{a}{1-r} + \frac{\lambda_i}{1-p}\right)\right)$$

Where INAR(1) defines a count process which has a marginal Poisson distribution with a modified parameter $P\left[\lambda_i \Theta_i / (1-p)\right]$

Therefore, the conditional distribution of Θ_i given $N_{i,-1}$ is

$$\gamma\left[a + N_{i,-1}, 1/\left(\frac{a}{1-r} + \frac{\lambda_i}{1-p}\right)\right]$$

The proof for other cases is similar to this and easy to generalized.

<u>Proposition 3.2</u> Prediction of the Course Impact Factor and (Frequency Part of) Pure Premium

(i) For $t = -1$, $E\left[\Theta_i\right] = 1 - r$, $P_{-1,i} = \lambda_i$

(ii) For $t = 0$, $E\left[\Theta_i \middle| N_{i,-1}\right] = \left(a + N_{i,-1}\right) / \left(\frac{a}{1-r} + \frac{\lambda_i}{1-p}\right)$

$$P_{0,i} = pN_{i,-1} + \lambda_i E\left[\Theta_i \middle| N_{i,-1}\right] = pN_{i,-1} + \left(a + N_{i,-1}\right) / \left(\frac{a}{\lambda_i (1-r)} + \frac{1}{1-p}\right)$$

(iii) For

$$t = 1, E\left[\Theta_i \middle| N_{i,-1}, N_{i,0}\right] = \frac{\sum_{z_2=0}^{\min(N_{i,-1},N_{i,0})} \pi\left(z_2, N_{i,-1}, N_{i,0}\right)\left(a + N_{i,-1} + N_{i,0} - z_2\right)}{\sum_{z_2=0}^{\min(N_{i,-1},N_{i,0})} \pi\left(z_2, N_{i,-1}, N_{i,0}\right)\left(\frac{a}{1-r} + \lambda_i + \frac{\lambda_i}{1-p}\right)}$$

$$P_{1,i} = pN_{i,0} + \lambda_i E\left[\Theta_i \middle| N_{i,-1}, N_{i,0}\right]$$

(iv) For $t = 2$,

$$E\left[\Theta_i \middle| N_{i,-1}, N_{i,0}, N_{i,1}\right] = \frac{\sum_{z_3=0}^{\min(N_{i,1},N_{i,0})} \sum_{z_2=0}^{\min(N_{i,-1},N_{i,0})} \pi\left(z_2, z_3, N_{i,-1}, N_{i,0}, N_{i,1}\right)\left(a + N_{i,-1} + N_{i,0} + N_{i,1} - z_2 - z_3\right)}{\sum_{z_3=0}^{\min(N_{i,1},N_{i,0})} \sum_{z_2=0}^{\min(N_{i,-1},N_{i,0})} \pi\left(z_2, z_3, N_{i,-1}, N_{i,0}, N_{i,1}\right)\left(\frac{a}{1-r} + 2\lambda_i + \frac{\lambda_i}{1-p}\right)}$$

$$P_{2,i} = pN_{i,1} + \lambda_i E\left[\Theta_i \middle| N_{i,-1}, N_{i,0}, N_{i,1}\right]$$

Proposition 3.3 Individual Discount Rate

(i) For $t = 0$, $r_{i,0} = \left(\dfrac{P_{0,i} - P_{-1,i}}{P_{-1,i}} \right)_+$

(ii) For $t = 1$, $r_{i,1} = \left(\dfrac{P_{1,i} - P_{0,i}}{P_{0,i}} \right)_+$

(iii) For $t = 2$, $r_{i,2} = \left(\dfrac{P_{2,i} - P_{1,i}}{P_{1,i}} \right)_+$

Where $r_{i,t}$ is the discount rate for the ith policy holder in the year t (t is a relative time to the training). $P_{-1,i}$, $P_{0,i}$, $P_{1,i}$ $P_{2,i}$ are estimated through Proposition 3.2.

3.4.2 Individual Discount Rate for BRC learners

For the BRC learners, we don't have the BRC learners' previous claim history record. One straightforward solution is to use the overall discount rate estimated in section 3.3, which already comprehensively considered the effectiveness of BRC program and other factors.

4 Numerical Illustrations

4.1 Estimation of Effectiveness

If we could cooperate with insurance company to obtain the individual level insurance claim data as well as the characteristics of those policy holders (to show whether they have taken any safety course or not), we could use the method in section 2 to estimate the effectiveness of BRC and ERC respectively. Because the data is unavailable, we will try to use the roughly estimated result from previous research to show our analytical framework. In James and Barbara(1989) "Does Motorcycle Training Reduce Crashes?" They showed overall the untrained group had 32 % more crashes than the trained group. John W Billheimer (1998) "Evaluation of California motorcyclist safety program" showed that analyses of statewide crashes trend indicate that fatal motorcycle crashes have dropped 69 percent since the introduction of the CMSP. While many other researchers showed that the program is not effective at all. For the illustration purposes, we assume the effectiveness of BRC is 10%, and the effectiveness for ERC is 6%.

4.2 Estimation of Overall Discount Rate for Motorcycle Safety Course

Table 5 CT total student registration

Year	BRC registration	ERC registration	Total Registered Motorcycles[9]

[9] Source: Highway Statistics http://www.fhwa.dot.gov/policy/ohim/hs03/htm/mv1.htm

2000	2680	229	54046
2001	3029	223	51883
2002	3203	222	62061
2003	4072	205	63071
2004	4363	205	64754
2005	5257	209	65110
2006	5588	139	64959
2007	5909	190	64848
2008	6315	150	65845
2009	4761	157	65900
2010	4676	106	66020
Total	49853	2035	
Percentage	96%	4%	

Based on the data in table 5, we could estimate $\alpha = 4\%, 1-\alpha = 96\%$, and r_E and r_B as 2% and 10% respectively. Using the rough value of r ($r = \partial r_E + (1-\partial) r_B$), the data e_{Et}, e_{Bt}, and N_t from 2000 to 2010, we estimate the values of parameter α and β in $e_{Et} = (N_t - e_{Bt})\left(1 - e^{-(\alpha r + \beta \rho_A)}\right)$ as 0.5 and 0.5 respectively.

In addition, since from January 2011, all applicants must successfully complete a novice safety course before obtaining a motorcycle endorsement. It can be anticipated that there would be a sharp rise on the 2011 BRC registration. Based on the pattern in table 5, let's suppose it return to the level in 2008, say 6300. While the number of total registered motorcycles seems to be relatively stable these years, let's forecast we have 66050 total registered motorcycles in 2011.

From the website of Rocky Mountain Insurance Information Association, we got the Cost of Auto Insurance by State[10] as follows. Here we assume the rate of motorcycle insurance is on the same level of other autos. Then base on the pattern in table 6, we estimate the average expenditures for 2009, 2010, and 2011 are 934, 920 and 915 respectively by the method of Two Moving Averages.

Table 6 CT average auto insurance expenditure

State	2008 Average Expenditure[11]	2008 Rank	2007 Average Expenditure	2007 Rank	2006 Average Expenditure	2006 Rank
Connecticut	$950	9	$964	10	$981	10

[10] Cost of auto insurance
http://www.rmiia.org/auto/steering_through_your_auto_policy/Cost_of_Auto_Insurance.asp

[11] The average insurance expenditure is calculated by adding all auto insurance premium collected for liability, comprehensive and collision coverage, and dividing by the number of insured cars for the year. This average is based on all policies - including liability-only and policies with optional comprehensive and collision coverage. Limits on policies vary widely and are based on state requirements as well as consumer choice.

Table 7 CT motorcycle crash number and involved person number

	Fatal Injury	Incapacitating injury	Non-incapacitating evident injury	Possible injury(claim of non-evident injury)	Not injured	Total Involved Person	Total Crash Number
Year	K	A	B	C	N		
1999	19	117	309	134	457	1036	995
2000	33	118	318	143	457	1069	1031
2001	32	115	366	184	497	1194	1154
2002	34	132	326	162	495	1149	1112
2003	19	122	316	155	502	1114	1069
2004	36	97	374	173	507	1187	1058
2005	28	136	398	171	575	1308	1266
2006	34	156	411	142	514	1257	1226
2007	26	194	489	187	639	1535	1480
2008	36	175	475	160	648	1494	1449

Based on the data in table 7, we get the estimated values of $L_{E,t-1}$ and $L_{N,t-1}$ as

$L_{E,t-1} = 103350$, $L_{N,t-1} = 60013240$.

Then $e_{Et} + e_{Nt} = N_t - e_{Bt} = 66050 - 6300 = 59750$

$$e_{Et} = \left(N_t - e_{Bt}\right)\left(1 - e^{-(\alpha r + \beta \rho_E)}\right) = 59750\left(1 - e^{-(0.5r+0.01)}\right)$$

$e_{E,t-1} = 106$, $e_{N,t-1} = 61238$, $P_t = 915$

Then the objective function turns into

$$MAX \quad 0.06 \times 59750\left(1 - e^{-(0.5r+0.01)}\right) \times \frac{103350}{106} + 0.1 \times 6300 \times \frac{60013240}{61238} - r\left(59750\left(1 - e^{-(0.5r+0.01)}\right) + 6300\right) \times 915$$

where $0 \le r \le 15\%$

It is easy to solve this nonlinear programming, and we get the value for r

$r = 0.0808 \approx 8\%$, and $\kappa = 0.835$

4.3 Estimation of Individual Discount Rate

The values of other parameters are chosen to be $p = 0.3$ and $\lambda_i = 0.3$ for comparison between different claim patterns. $r = 8\%$, $E\left[\Theta_i\right] = 1 - r$. Then using the result in Proposition 3.1 to 3.3, we could obtain the estimated result in table 8 to table 10.

Table 8 Expected value of the training impact factor given the claim history

Claim History	$E\left[\Theta_i \middle\vert N_{i,-1}\right]$	$E\left[\Theta_i \middle\vert N_{i,-1}, N_{i,0}\right]$	$E\left[\Theta_i \middle\vert N_{i,-1}, N_{i,0}, N_{i,1}\right]$
(0,0,1)	0.768496	0.689065	0.936773
(0,1,0)	0.768496	1.033597	0.936773

(1,2,0)	1.152745	1.654619	1.344014
(1,0,1)	1.152745	1.033597	1.24903
(1,1,0)	1.152745	1.329571	1.148302
(0,1,1)	0.768496	1.033597	1.12261
(2,0,0)	1.536993	1.37813	1.24903
(2,1,0)	1.536993	1.654619	1.344014
(3,0,0)	1.921241	1.722662	1.561288
(1,0,0)	1.152745	1.033597	0.936773

Table 9 Prediction of (the frequency part of) Pure Premium

Claim History	$P_{-1,i}$	$P_{0,i}$	$P_{1,i}$	$P_{2,i}$
(0,0,1)	0.3	0.230549	0.20672	0.581032
(0,1,0)	0.3	0.230549	0.610079	0.281032
(1,2,0)	0.3	0.645823	1.096386	0.403204
(1,0,1)	0.3	0.645823	0.310079	0.674709
(1,1,0)	0.3	0.645823	0.698871	0.344491
(0,1,1)	0.3	0.230549	0.610079	0.636783
(2,0,0)	0.3	1.061098	0.413439	0.374709
(2,1,0)	0.3	1.061098	0.796386	0.403204
(3,0,0)	0.3	1.476372	0.516799	0.468386
(1,0,0)	0.3	0.645823	0.310079	0.281032

Table 10 Individual Training Discount for Motorcycle Insurance

Claim History	$r_{0,i}$	$r_{1,i}$	$r_{2,i}$
(0,0,1)	0.231504	0.10336	0
(0,1,0)	0.231504	0	0.539352
(1,2,0)	0	0	0.632242
(1,0,1)	0	0.51987	0
(1,1,0)	0	0	0.507076
(0,1,1)	0.231504	0	0
(2,0,0)	0	0.610367	0.093678
(2,1,0)	0	0.24947	0.493707
(3,0,0)	0	0.649954	0.093677
(1,0,0)	0	0.51987	0.093677

We should be notice that there may be other factor would affect the claim history of customers, not just because of the training. Therefore, for those values of discount rate higher than 10%, we will use 10%, if lower than 10%, we will use the estimated value. $\overline{r}_{t,i} = \min(r_{t,i}, 10\%)$

5 Conclusion

The genesis of this paper comes from the debates about introducing of Mandatory Motorcycle Safety Course in Connecticut in January 2011. However, most research or projects which focus on evaluating the effectiveness of motorcycle safety course systematically and comprehensively are before 2000, and have shown mixed results on effectiveness. We also noticed that the safety courses are uniformly designed and guided by MSF, but the discount rate offered by different insurance companies varies greatly. Therefore, this paper has developed a unifying framework to quantify the effectiveness of such mandatory programs and to translate this in terms of a possible insurance rate reduction. First, the definition of effectiveness is given and formulated for both Basic and Advanced Motorcycle Safety Course by the measure of reduction in incurred loss or claim frequency. Then, evaluation about current most widely adopted "3-year" constraints is carried out.

Next, our research is divided into two major steps: Overall Discount Rate Estimation and Individual Discount Rate Adjustment by Personal Claim History. For the Overall Discount Rate, we use past insurance claim data to estimate the overall discount rate in nonlinear optimization programming, taken into consideration both the training effect on pure premium and the demand. For the Individual Discount Rate Adjustment, the Integer-Valued Autoregressive (INAR) method is used to model individual annual claim count in consecutive policy years and interpret the heterogeneity Θ_i as training impact factor. Through the strict theoretical derivation, conditional distribution of Θ_i, prediction of the Course Impact Factor and (Frequency Part of) Pure Premium, as well as Individual Discount Rate are derived.

To illustrate our analytical framework, numerical Illustrations are given in the last part of this party. Since the individual level insurance claim data as well as the characteristics of those policy holders (to show whether they took the safety training or not) are temporarily unavailable, we made some assumptions about the effectiveness for BRC and ARC respectively; but registered motorcycles, Average Auto Insurance Expenditure, Motorcycle Crash Number and Involved Person Number are based on Connecticut data. Based on our assumption, the overall discount rate for motorcycle safety course should be 8%, which may not reflect the real situation but can serve as an illustration. If insurance claim data were available, we could update the assumptions in our model and determine the estimation for overall discount rate. For the Individual Discount Rate stage, we considered 10 different kinds typical claim histories in the past three years and generated the Individual Training Discount Adjustment table. Possible future work could be the collection and combination of the data both from the transportation, health department, and insurers, to conduct the analysis presented in this paper; or introduce more stochastic factors in the calibration of effectiveness.

6 References

[1] Al-Osh, M. A. and A. A. Alzaid (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis 8*, 261-275.

[2] Billheimer, J. Evaluation of the California motorcyclist safety program.(1998) Presented at the annual meeting of the Transportation Research Board, Washington, DC

[3] Deutermann, W. 2004. Motorcycle helmet effectiveness revisited. Report no. DOT HS-809-715. Washington, DC: National Highway Traffic Safety Administration.

[4] Gourieroux, C. and J. Jasiak (2004). Heterogeneous INAR(1) Model with Application to Car Insurance. *Insurance: Mathematics and Economics 34*, 177-192.

[5] Highway Loss Data Institute Bulletin. December 2009. Motorcycle collision coverage claims in states with required motorcycle rider training. Vol 26(12).

[6] James and Barbara(1989),Does Motorcycle Training Reduce Accidents? Evidence from a Longitudinal. Quasi-experimental Study. Journal of Safety Research.20: 61-72

[7] Jonah, B, Davidson, N., Bragg B. Are formally trained motorcyclists safer" Accident Analysis and Prevention, Vol 14(4), 1982.

[8] John W Billheimer (1998) Evaluation of California motorcyclist safety program. Transportation Research Record, Paper No.98-0652

[9] Leigh J. Halliwell, (2005), "Utility-Theoretic Underwriting," http://www.casact.org/sections/care/0905/handouts/halliwell.pdf

[10] McDavid, J., Lohrmann, B, Lohrmann, G. 1989. Does motorcycle training reduce accidents? Evidence for a longitudinal quasi-experimental study. J Safety Res. Vol 20, p 61-72.

[11] McKnight, J. Evaluation of the Pennsylvania motorcycle safety program, Final report, prepared for the Indiana University of Pennsylvania by the National Public service Research Institute., Landover, MD, 1987.

[12] National Highway Traffic Safety Administration. 2009. Traffic safety facts, 2008. Washington, DC: US Department of Transportation.

[13] Pacific Institute for Research and Evaluation, (2010 updated) Injury Prevention: What Works? A summary of Cost-Outcome Analysis for Motor vehicle

[14] Ting Zhang, (2009) Integer-Valued Autoregressive Processes with Dynamic Heterogeneity and their Applications in Automobile Insurance, Master Degree Thesis,

[15] Tim Buche, (2010),Giving Motorcyclists The Best in Training: Designing Principle-Based, Safety-Oriented Education And Training Programs, Motorcycle Safety Foundation, 2010 Conference

[16] Tim Buche, Sherry Williams, and Ray Ochs,(2010),Motorcycle Safety Foundation, MSF RETS: A System Designed To Succeed, presented in support of the Vulnerable Road Users Conference in Jerusalem, May 30-June 2, 2010.

[17] Walhin, J. F. and J. Paris (2000). The true claim amount and frequency distributions within a bonus-malus system. *ASTIN Bulletin 30*, 391-403.