Preliminary Results

José I. Velarde Morales

August 11, 2022

1 Summary

- I implemented the one zone model for both the minimum premium model and the minimum CVaR model.
- The minimum CVaR model seems to give more reasonable payout functions, but in both cases the full model, ie the model that uses both an intercept and a slope for the payout function, always has a payout of 0.
- I tested a range of parameter values for both models to try to get a better sense of what's going on.

2 Setup

Data Generating Process

We use the following data generating processes for the toy examples. In order to make the comparison with the status quo more straightforward, we model losses instead of overall wealth.

$$\begin{aligned}
&l = \beta\theta + \epsilon \\
&\theta \sim \mathcal{N}(5, 2) \\
&\epsilon \sim \mathcal{N}(0, \sigma I) \\
&\beta = 3
\end{aligned}$$

I draw 100 training samples from the above model to train the prediction model and to use as input for the optimization programs. I then evaluate both methods using 1000 samples drawn from the same model.

Notes/Things to improve

- I also want to try this toy example using a model that generates mortality rates instead of levels, however, my first attempts at this failed.
- I also need to think more carefully about what distribution to draw from for this toy example. As it stands, losses are always positive, which is unrealistic.
- Also want to try $l = f(\theta) + \epsilon$ where $\theta \sim \mathcal{N}(\mu, \Sigma)$, f is nonlinear, and $\epsilon \sim \mathcal{N}(0, \sigma I)$.

Performance Metrics

We will be using the following performance metrics to compare the two approaches. Note, in this document we will only be presenting the results for the one with $*$.

- Probability of farmer ruin: probability that loss net of insurance exceeds a pre-specified threhsold. $*$
- Premium$*$
- Accuracy: percentage of correct decisions (i.e. giving a payout when a covered loss occured, and not giving a payout when a covered loss does not occur.)
- Insurer Basis Risk: probability that the contract will pay out when a covered loss did not occur.
- Insured Party Basis Risk: probability that the contract will not pay out when a covered loss occurs.
- Share of losses covered
- Average cost to insurer, should also include extreme cases.

Note: We should figure out how we will weigh all these different metrics, and we should review the metrics mentioned in Risk Modeling for WII.

3 Status Quo

We will be comparing our proposed approach to the method developed by Chantarat et al. (2013), which, to the extent of our knowledge is what is currently being used for Kenya's Index Based Livestock Insurance (IBLI) program. The method is as follows:

- Use a clustering algorithm to group locations into clusters
- Fit a separate herd mortality function for each cluster. They use a linear regression model to predict herd mortality.
- Contracts are of the form: $I(\theta) = \max(\hat{M}(\theta) - M^*, 0) \times TLU \times P_{TLU}$ where TLU is the number of insured livestock units, and P_{TLU} is the price per insured livestock unit. In other words, their contract pays farmers for the full predicted loss beyond a threshold, M^*. This threshold, M^* is the contract's strike value.
- They choose the strike value that would explain the highest share of insurable losses in the historical data. In other words, they run the following regression: $y_s = \beta_s \hat{y_s} + \epsilon$ where y_s is the actual insured losses at strike value s and $\hat{y_s}$ is the predicted insured losses at strike value s. They choose the strike value $s = \arg\max_s \beta_s$.

To mimick this in our toy example, we set the status quo contracts to be $I(\theta) = \max(\hat{l}(\theta) - l^*, 0)$, since we are already assuming that l is the total loss suffered. For the toy example, we fit a (correctly specified) linear regression model to predict losses: $l = \beta\theta + \epsilon \implies \hat{l}(\theta) = \hat{\beta}\theta$.

4 Optimization Approach

In order to separate the effect of contract design from the effect of prediction quality, we will be basing our contracts on the same predictions used by the status quo method. In other words, we will use the status quo method to estimate a model that predicts loss based on theta, $\hat{l}(\theta)$, and our payout function will use that as input instead of θ. In other words, our model will define payout functions $I(\hat{l}(\theta))$, where $\hat{l}(\theta)$ is the same prediction function used by the status quo method.

Minimum CVaR Model

This model minimizes the $CVaR$ of the farmer's net loss subject to a constraint on the premium. The premium constraints are expressed as a fraction of the full insured amount.

Single Zone Model

$$\min_{a,b,\pi} \quad t + \frac{1}{\epsilon}\sum_k p_k \gamma_k \tag{1}$$

$$\text{s.t. } I^k \geq a\hat{l}(\theta^k) + b, \forall k \tag{2}$$

$$0 \leq I^k \leq y, \forall k \tag{3}$$

$$E[I] \leq \bar{\pi} y \tag{4}$$

$$\gamma_k \geq l + E[I] - I^k - t, \forall k \tag{5}$$

$$\gamma_k \geq 0, \forall k \tag{6}$$

Here, y is the maximum insured amount, and k indexes the possible realizations of θ, l.

Model Parameters

- ϵ: This defines the CVaR objective. $\epsilon = 0.1$ means that our objective is on the expected value of the loss given that it is above the 90^{th} percentile.
- $\bar{\pi}$: This is the maximum value of the premium.

Minimum Premium Model

This model minimizes the premium subject to a maximum $CVaR$ constraint. This can be interpreted either as an approximation of the original probability constraint, or as a constraint on the expected value of the loss beyond a given threshold.

Single Zone Model

$$\min_{a,b,\pi} E[I^k] \tag{7}$$

$$\text{s.t. } I^k \geq a\hat{l}(\theta^k) + b, \forall k \tag{8}$$

$$0 \leq I^k \leq y, \forall k \tag{9}$$

$$t + \frac{1}{\epsilon}\sum_k p_k \gamma_k \leq \bar{l} \tag{10}$$

$$\gamma_k \geq l^k + E[I] - I^k - t, \forall k \tag{11}$$

$$\gamma_k \geq 0, \forall k \tag{12}$$

Here, y is the maximum insured amount, $\bar{l}$ is the upper constraint on the $CVaR$ of the loss, and k indexes the possible realizations of θ, l.

Model Parameters

- ϵ: This defines the tail for the CVaR constraint. $\epsilon = 0.1$ means that our CVaR constraint is on the expected value of the loss given that it is above the 90^{th} percentile.
- $\bar{l}$: This is the bound for the CVaR constraint.

5 Results

The model didn't behave as I expected, so I ran some experiments to try to get a better understanding. I tried three different versions of each model: a full model that fits both slope and intercept ($I(\theta) = a\theta + b$), an intercept only model ($I(\theta) = \theta + b$), and a slope only model ($I(\theta) = a\theta$). For each version, I varied the parameters, $\epsilon, \bar{l}, \bar{pi}$ depending on the moel. I also ran a version of the CVaR model with the additional constraint that $a \geq 0.5$, and it gave non-zero payout functions.

CVaR Model

Full Model

Figure 1: Payout Functions

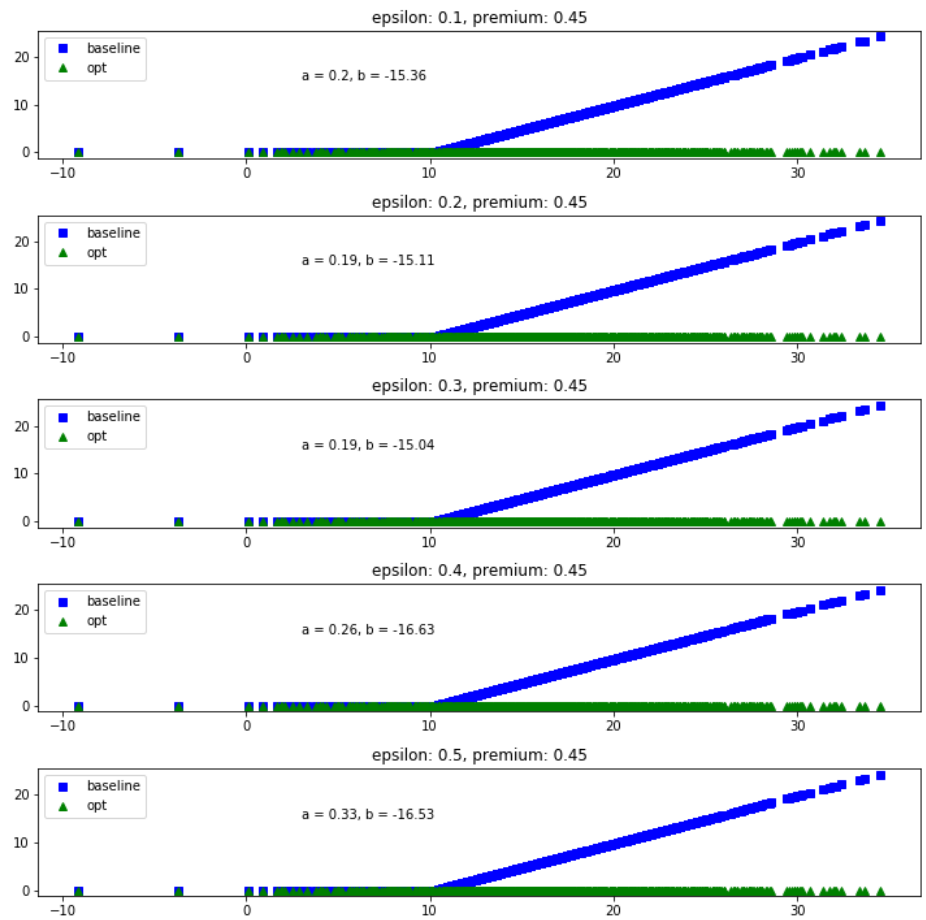

Table 1: Performance Metrics

	Eps	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.1	0.30	0.20	0.10	0.05	0.00
2	0.2	0.30	0.20	0.10	0.05	0.00
4	0.3	0.30	0.20	0.10	0.05	0.00
6	0.4	0.30	0.20	0.10	0.05	0.00
8	0.5	0.30	0.20	0.10	0.05	0.00
1	Baseline	0.00	0.00	0.00	0.00	7.02

Epsilon Exploration

Figure 2: Payout Functions

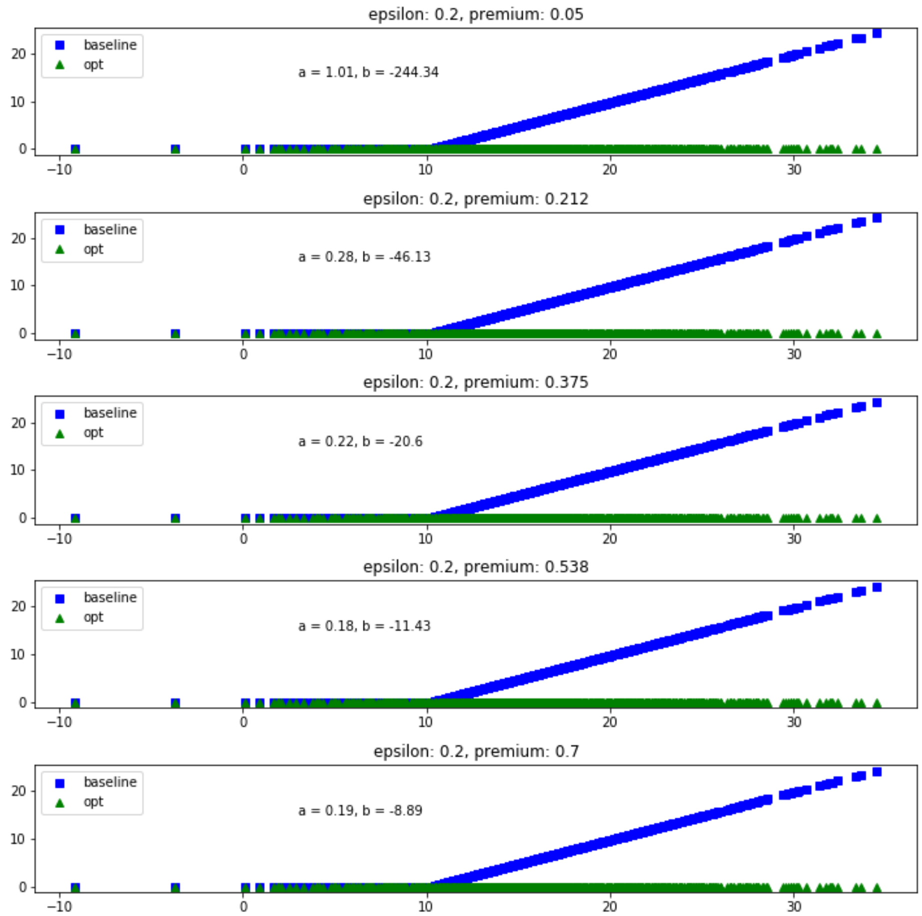

Table 2: Performance Metrics

	Max Premium	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.05	0.30	0.20	0.10	0.05	0.00
2	0.212	0.30	0.20	0.10	0.05	0.00
4	0.375	0.30	0.20	0.10	0.05	0.00
6	0.538	0.30	0.20	0.10	0.05	0.00
8	0.7	0.30	0.20	0.10	0.05	0.00
1	Baseline	0.00	0.00	0.00	0.00	7.02

Premium Exploration

Intercept Only Model

Figure 3: Payout Functions

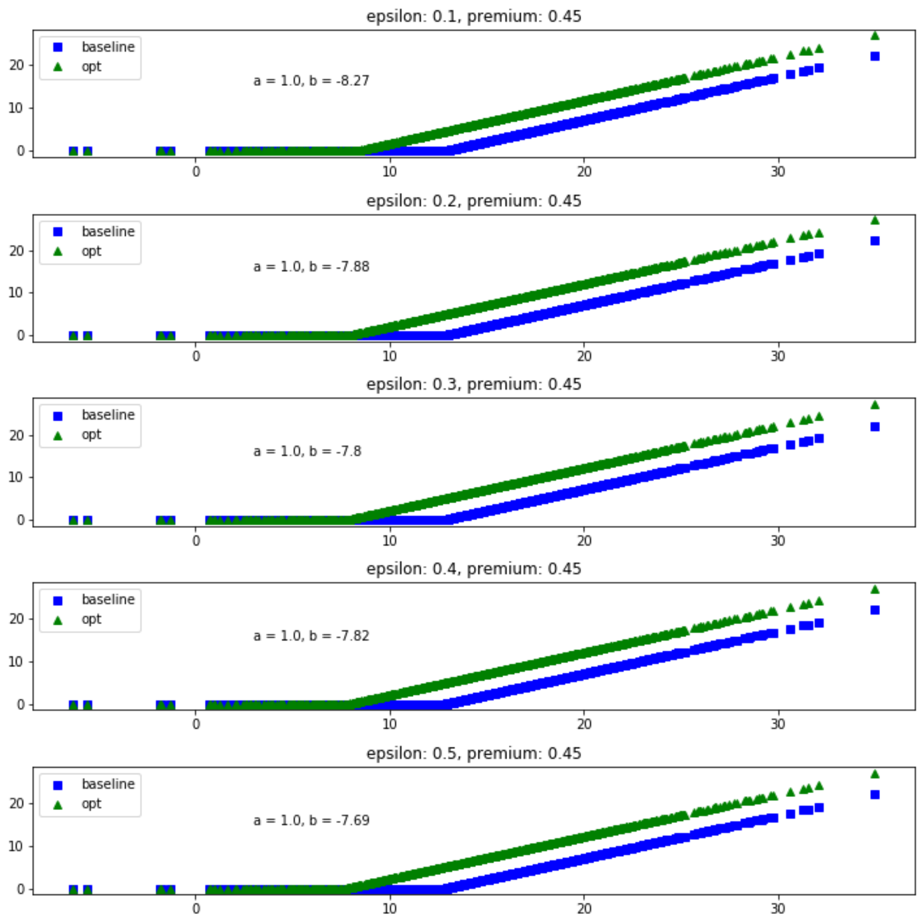

Table 3: Performance Metrics

	Eps	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.1	0.00	0.00	0.00	0.00	7.09
2	0.2	0.00	0.00	0.00	0.00	7.44
4	0.3	0.00	0.00	0.00	0.00	7.51
6	0.4	0.00	0.00	0.00	0.00	7.49
8	0.5	0.00	0.00	0.00	0.00	7.61
1	Baseline	0.00	0.00	0.00	0.00	3.55

Epsilon Exploration

Figure 4: Payout Functions

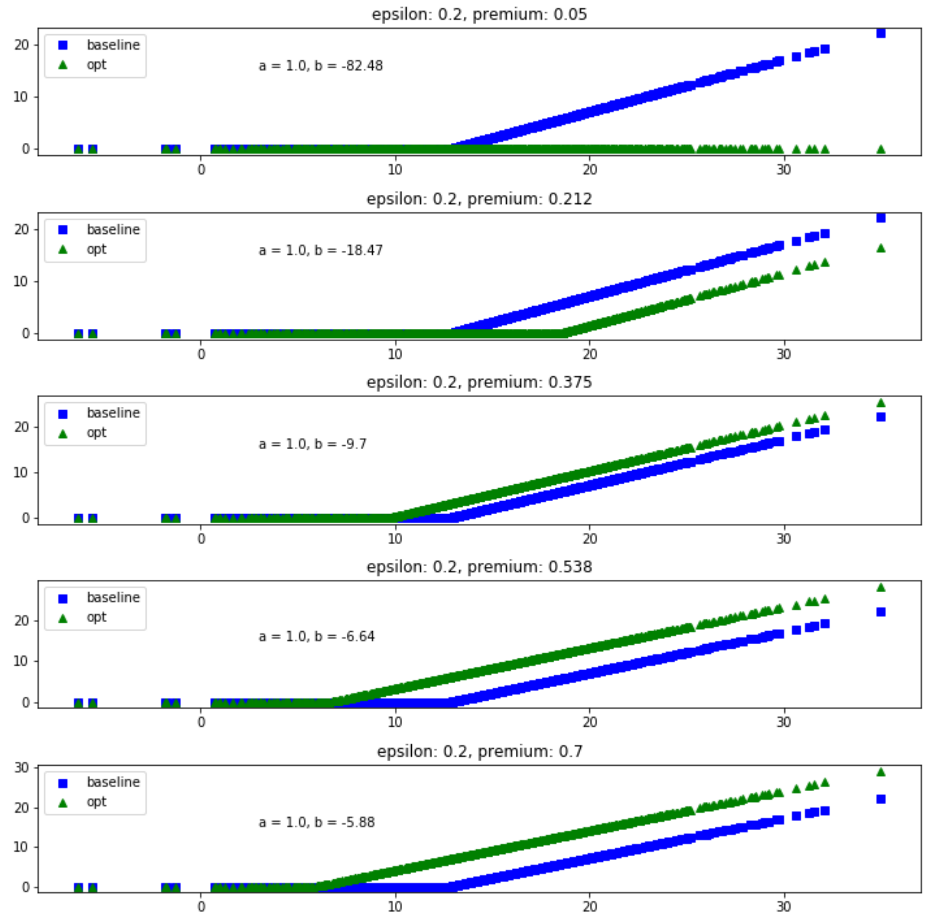

Table 4: Performance Metrics

	Max Premium	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.05	0.30	0.20	0.10	0.05	0.00
2	0.212	0.30	0.00	0.00	0.00	1.05
4	0.375	0.00	0.00	0.00	0.00	5.88
6	0.538	0.00	0.00	0.00	0.00	8.56
8	0.7	0.00	0.00	0.00	0.00	9.27
1	Baseline	0.00	0.00	0.00	0.00	3.55

Premium Exploration

Slope Only Model

Figure 5: Payout Functions

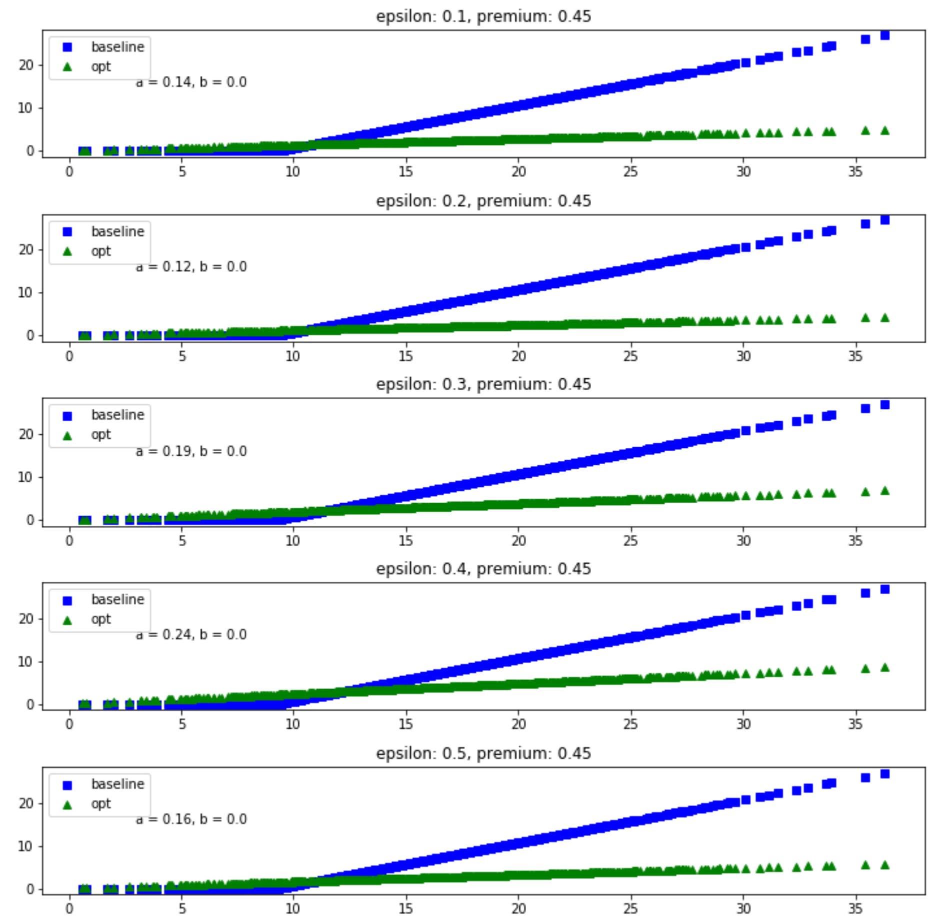

Table 5: Performance Metrics

	Eps	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.1	0.14	0.07	0.02	0.01	2.37
2	0.2	0.16	0.09	0.03	0.01	2.03
4	0.3	0.09	0.04	0.01	0.01	3.22
6	0.4	0.05	0.02	0.01	0.00	4.06
8	0.5	0.12	0.06	0.01	0.01	2.71
1	Baseline	0.00	0.00	0.00	0.00	7.83

Epsilon Exploration

Figure 6: Payout Functions

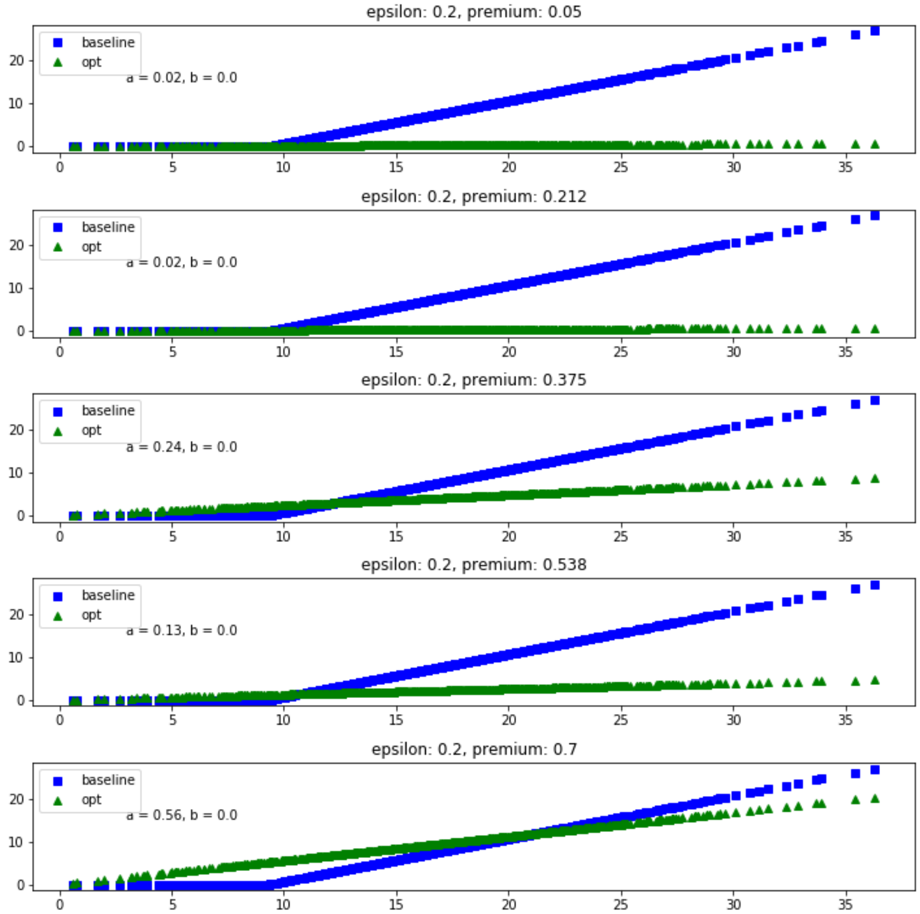

Table 6: Performance Metrics

	Max Premium	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.05	0.27	0.18	0.08	0.04	0.34
2	0.212	0.27	0.18	0.08	0.04	0.34
4	0.375	0.05	0.02	0.01	0.00	4.06
6	0.538	0.15	0.08	0.02	0.01	2.20
8	0.7	0.00	0.00	0.00	0.00	9.48
1	Baseline	0.00	0.00	0.00	0.00	7.83

Premium Exploration

Full Model, $a \geq 0.5$

Figure 7: Payout Functions

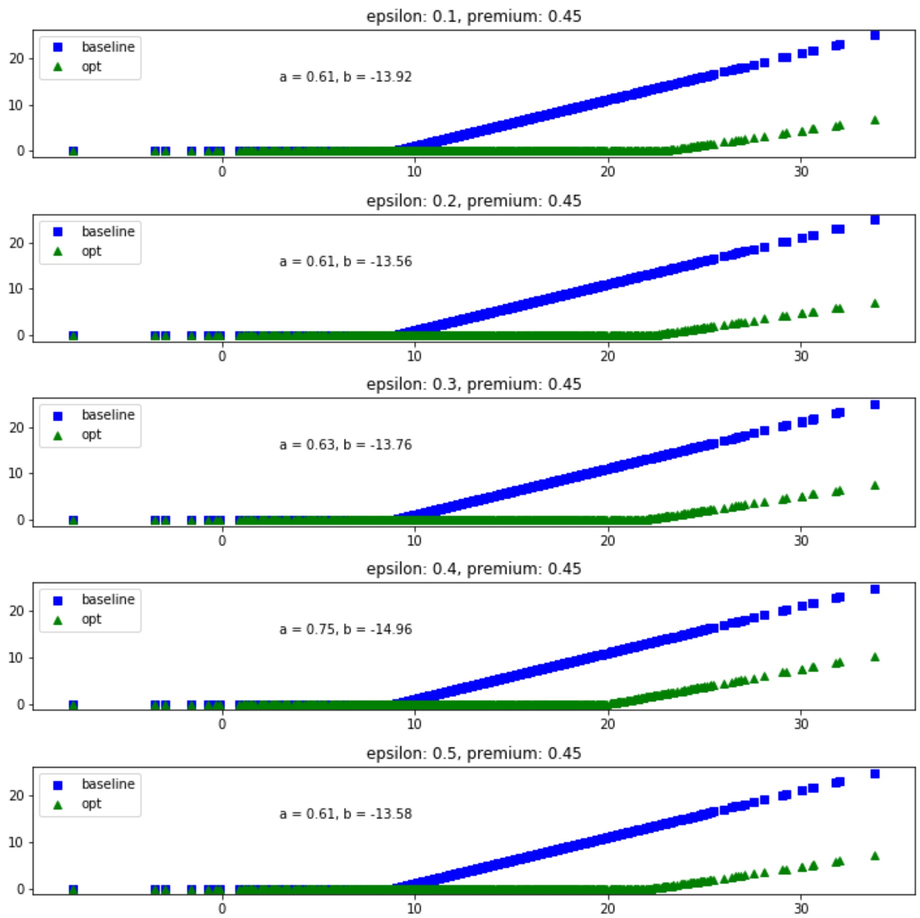

Table 7: Performance Metrics

	Eps	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.1	0.30	0.20	0.10	0.04	0.09
2	0.2	0.30	0.20	0.10	0.04	0.11
4	0.3	0.30	0.20	0.10	0.02	0.14
6	0.4	0.30	0.20	0.03	0.00	0.31
8	0.5	0.30	0.20	0.10	0.04	0.11
1	Baseline	0.00	0.00	0.00	0.00	5.16

Epsilon Exploration

Figure 8: Payout Functions

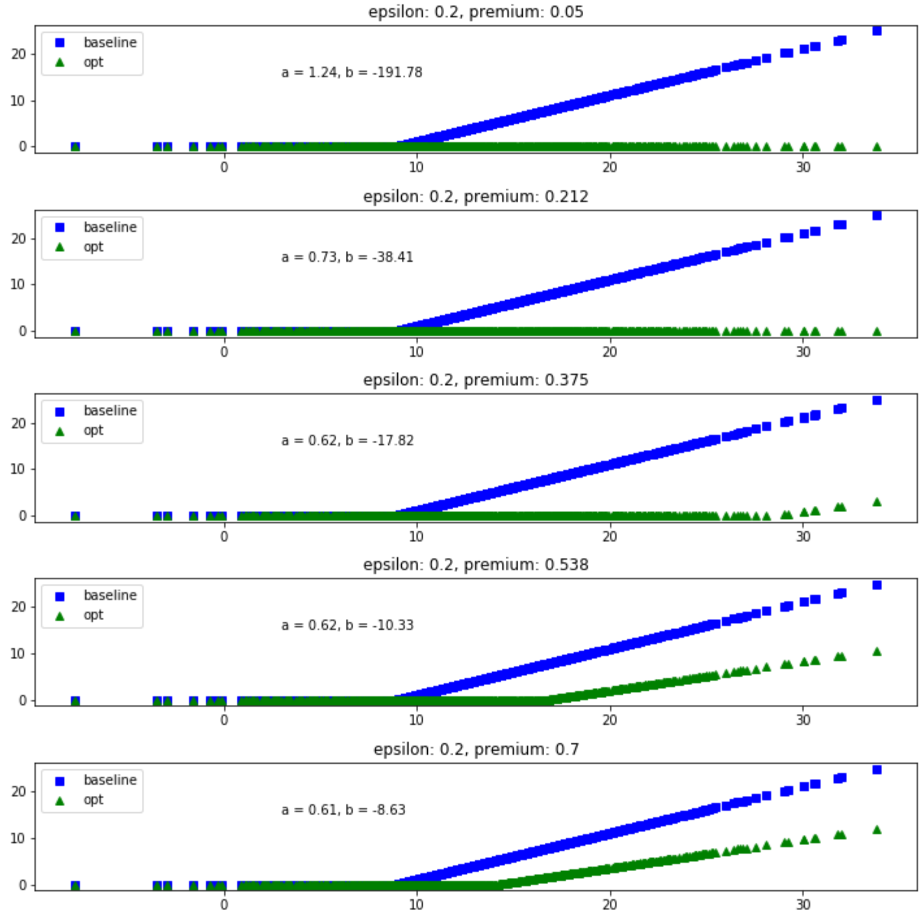

Table 8: Performance Metrics

	Max Premium	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.05	0.30	0.20	0.10	0.05	0.00
2	0.212	0.30	0.20	0.10	0.05	0.00
4	0.375	0.30	0.20	0.10	0.05	0.01
6	0.538	0.30	0.11	0.01	0.00	0.67
8	0.7	0.13	0.03	0.00	0.00	1.21
1	Baseline	0.00	0.00	0.00	0.00	5.16

Premium Exploration

Premium Model

Full Model

Figure 9: Payout Functions

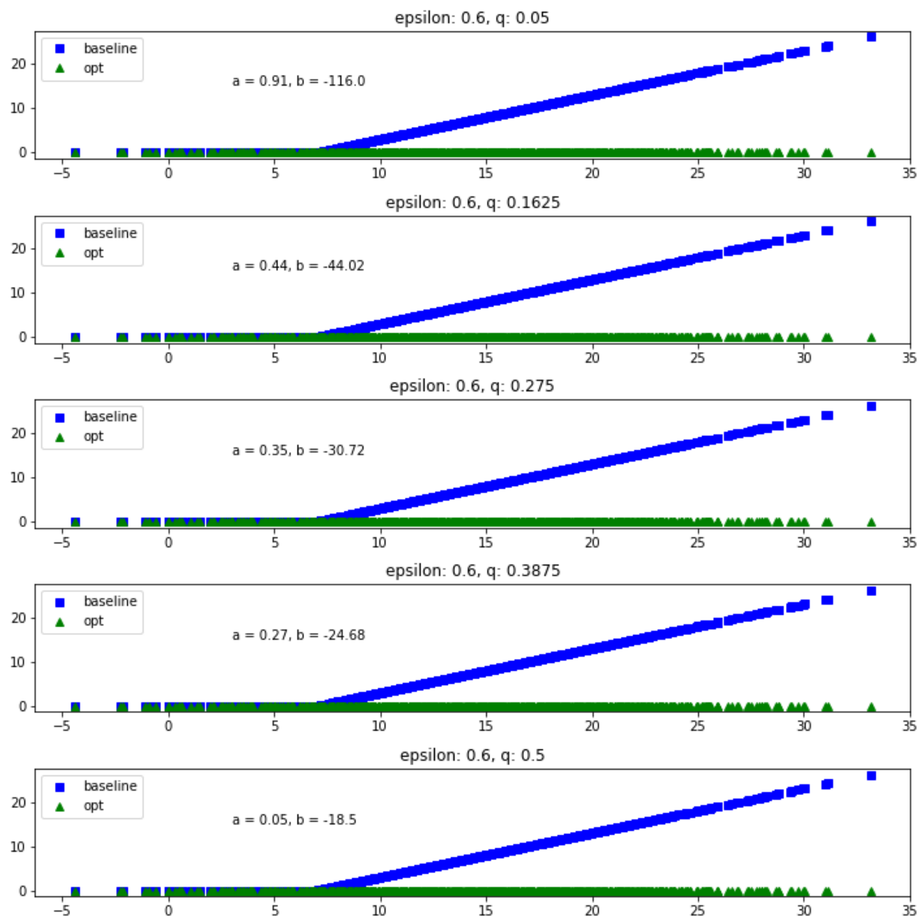

Table 9: Performance Metrics

	Eps	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.05	0.30	0.20	0.10	0.05	0.00
2	0.1625	0.30	0.20	0.10	0.05	0.00
4	0.275	0.30	0.20	0.10	0.05	0.00
6	0.3875	0.30	0.20	0.10	0.05	0.00
8	0.5	0.30	0.20	0.10	0.05	0.00
1	Baseline	0.00	0.00	0.00	0.00	8.16

Epsilon Exploration

Figure 10: Payout Functions

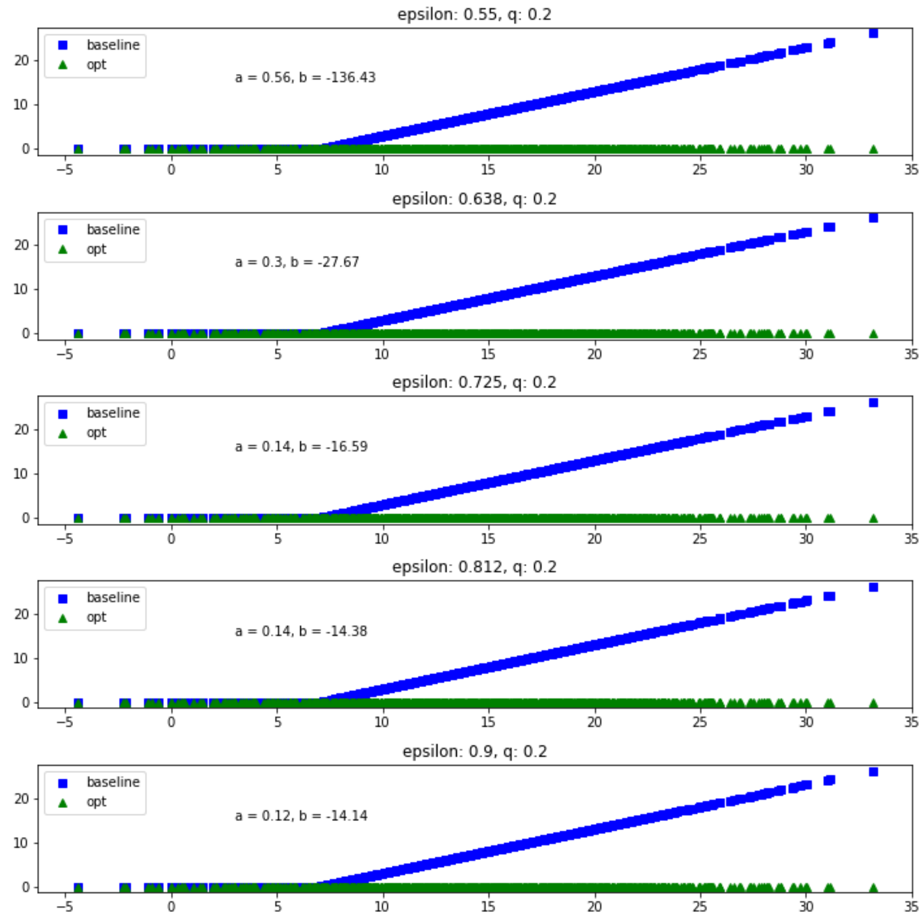

Table 10: Performance Metrics

	Loss q	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.55	0.30	0.20	0.10	0.05	0.00
2	0.638	0.30	0.20	0.10	0.05	0.00
4	0.725	0.30	0.20	0.10	0.05	0.00
6	0.812	0.30	0.20	0.10	0.05	0.00
8	0.9	0.30	0.20	0.10	0.05	0.00
1	Baseline	0.00	0.00	0.00	0.00	8.16

Premium Exploration

Intercept Only Model

Figure 11: Payout Functions

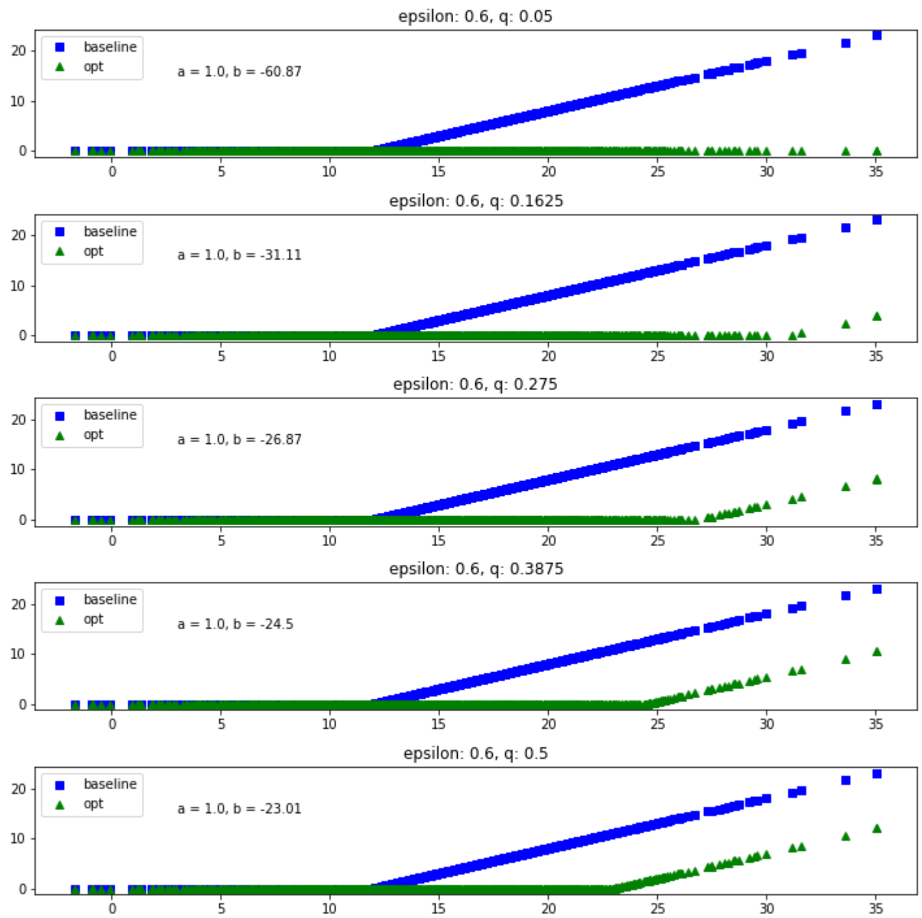

Table 11: Performance Metrics

	Eps	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.05	0.30	0.20	0.10	0.05	0.00
2	0.1625	0.30	0.20	0.10	0.05	0.01
4	0.275	0.30	0.20	0.10	0.05	0.05
6	0.3875	0.30	0.20	0.10	0.00	0.12
8	0.5	0.30	0.20	0.10	0.00	0.23
1	Baseline	0.00	0.00	0.00	0.00	4.26

Epsilon Exploration

Figure 12: Payout Functions

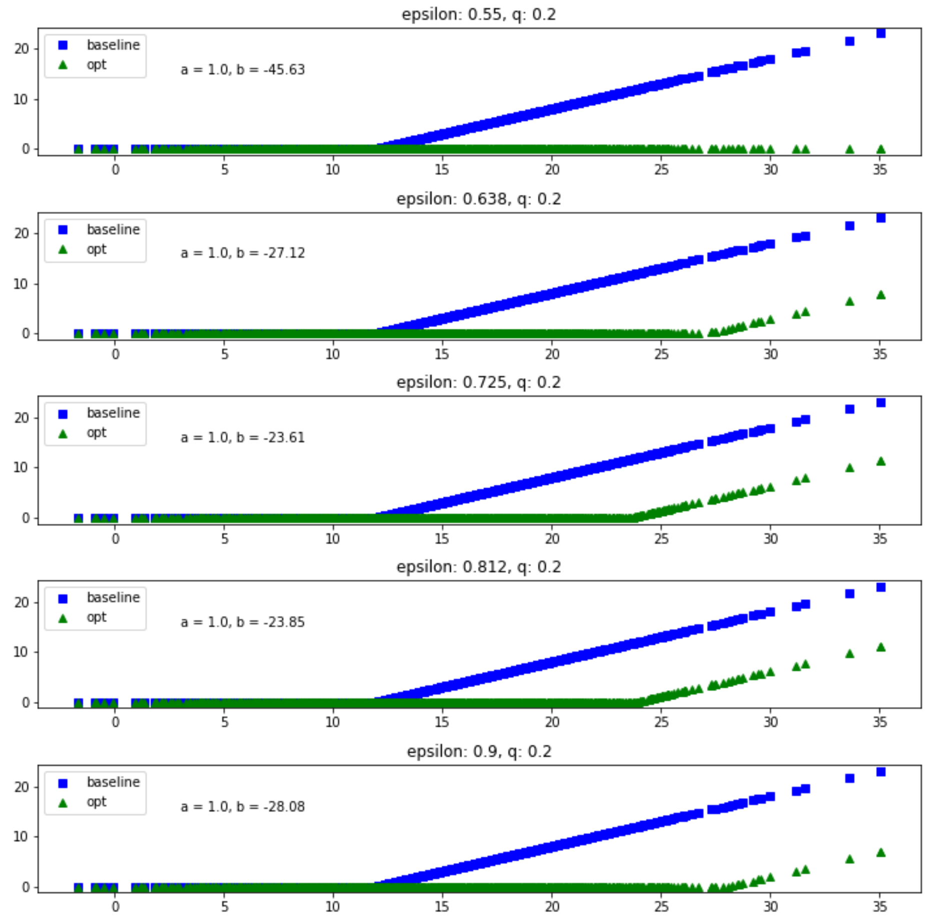

Table 12: Performance Metrics

	Loss q	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.55	0.30	0.20	0.10	0.05	0.00
2	0.638	0.30	0.20	0.10	0.05	0.05
4	0.725	0.30	0.20	0.10	0.00	0.18
6	0.812	0.30	0.20	0.10	0.00	0.16
8	0.9	0.30	0.20	0.10	0.05	0.03
1	Baseline	0.00	0.00	0.00	0.00	4.26

Loss Threshold Exploration

Slope Only Model

Figure 13: Payout Functions

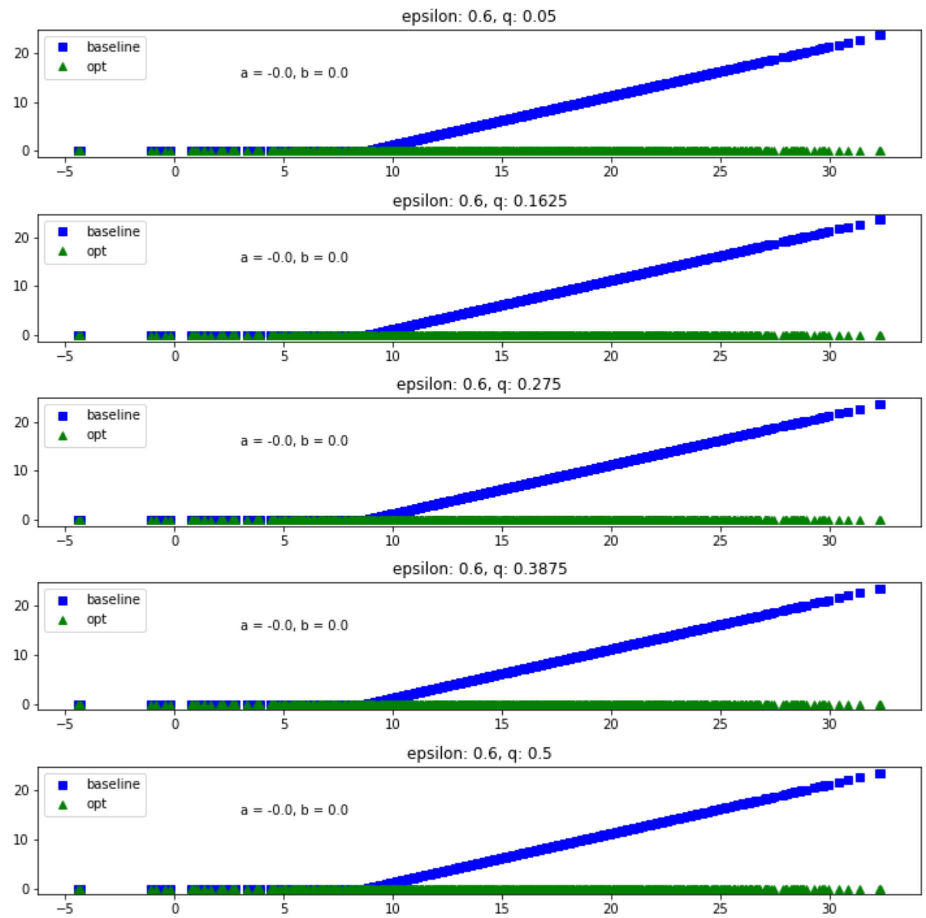

Table 13: Performance Metrics

	Eps	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.05	0.30	0.20	0.10	0.05	0.00
2	0.1625	0.30	0.20	0.10	0.05	0.00
4	0.275	0.30	0.20	0.10	0.05	0.00
6	0.3875	0.30	0.20	0.10	0.05	0.00
8	0.5	0.30	0.20	0.10	0.05	0.00
1	Baseline	0.00	0.00	0.00	0.00	7.98

Epsilon Exploration

Figure 14: Payout Functions

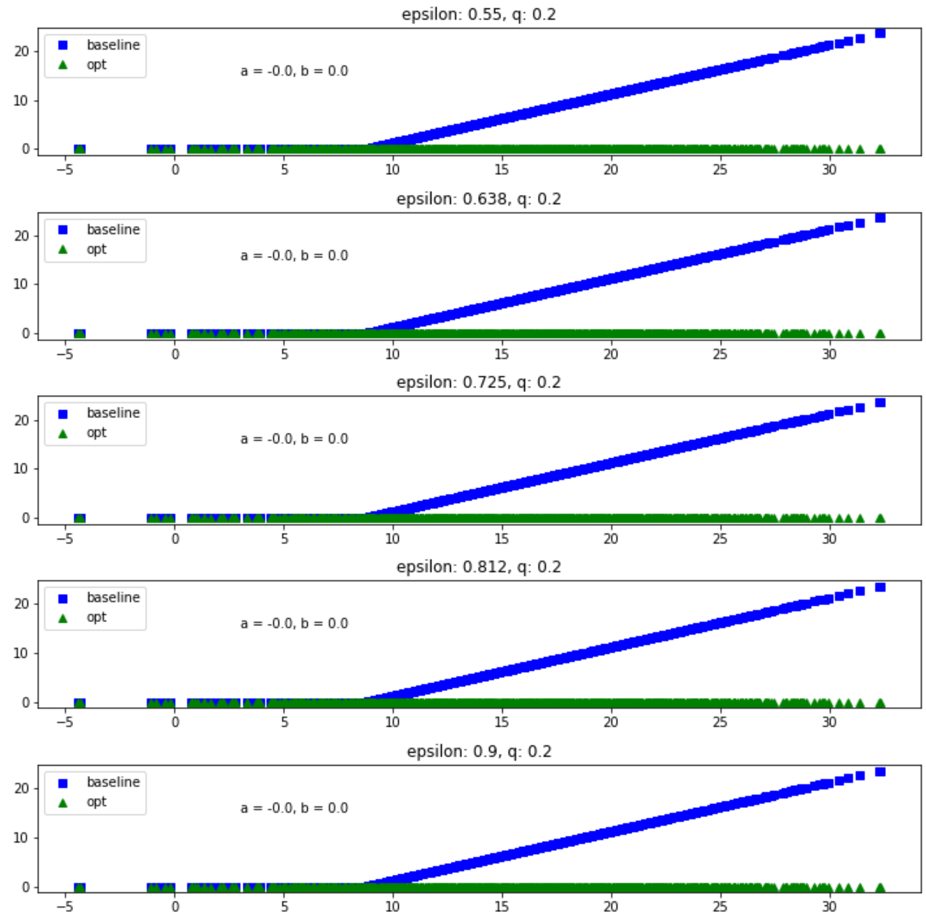

Table 14: Performance Metrics

	Loss q	P(L ¿ 70)	P(L ¿ 80)	P(L ¿ 90)	P(L ¿ 95)	Premium
0	0.55	0.30	0.20	0.10	0.05	0.00
2	0.638	0.30	0.20	0.10	0.05	0.00
4	0.725	0.30	0.20	0.10	0.05	0.00
6	0.812	0.30	0.20	0.10	0.05	0.00
8	0.9	0.30	0.20	0.10	0.05	0.00
1	Baseline	0.00	0.00	0.00	0.00	7.98

Loss Threshold Exploration