MACS 30200

Methods and Results

Xi Chen

Data and Methods

This paper uses the data from the experiment conducted by Ericson et al., (2015), which recruited 940 participants from the Amazon Mechanical Turk. The experiment included five conditions, each of which included a MEL task. The MEL tasks in different conditions varied in the representations of money terms (absolute or relative) and the framing of the time terms (delay or speedup): [1]

- Condition 1: Absolute Money Value, Delay Framing

  (e.g., $5 today vs. $5 plus an additional $5 in 4 weeks)

- Condition 2: Relative Money Value, Delay Framing

  (e.g., $5 today vs. $5 plus an additional 100% in 4 weeks)

- Condition 3: Standard MEL Format

  (e.g., $5 today vs $10 in 4 weeks)

- Condition 4: Absolute Money Value, Speedup Framing

  (e.g., $10 in 4 weeks vs. $10 minus $5 today)

- Condition 5: Relative Money Value, Speedup Framing

  (e.g., $10 in 4 weeks vs. $10 minus 50% today)

[1] In the following analysis, "Condition 0" means the pooled data including all observations from the five conditions.

Each participant was randomly assigned to one of the above five conditions and asked to answer 25 questions, each of which had varied money and time amounts: the money values range from \$0.01 to 100,000.00 and the time amounts range from 0 weeks to 6 weeks.

Table 1 presents descriptive statistics for the five key variables: *X1* refers to the monetary value associated with the smaller, sooner option; *T1* refers to the time amount associated with the smaller, sooner option; *X2* refers to the monetary value associated with the larger, later option; *T2* refers to the time amount associated with the larger, later option; *LateroptionChosen* refers to the option the participant chose: 1 if the larger, later option was chosen, and 0 if not.

Table 1. Summary Statistics for Key Variables

	X1	T1	X2	T2	LaterOptionChosen
mean	2930.618	1.002	3098.422	3.007	0.372
std	12756.33	0.8143	12895.23	1.155	0.483
min	0.03	0	0.13	1	0
25%	2.5	0	5.5	2	0
50%	40	1	55	3	0
75%	500	2	1000	4	1
max	100000	2	101000	5	1

To compare models, the main method used in this paper is to perform cross-validation analyses which split the data into two randomly generated parts: 75% for training and 25% for testing. In the intertemporal choice research, the most common used measure of model fit is the Mean Absolute Deviation (MAD) between a model's prediction and the actual choice, both of which are probabilities of choosing the larger, later option. To estimate the parameters of the models, maximum likelihood estimation is performed.

Results

Comparing Model Fit Across Six Models

To begin with, a model fit comparison is conducted across six typical models: Exponential, Hyperbolic, Quasi-hyperbolic, Tradeoff, DRIFT, and ITCH model. As Figure 1 shows, consistent with the results presented in Ericson et al., (2015), the three heuristic models outperform the three delay discounting models, and the ITCH model has the best performance. Figure 2 presents all models' performance across conditions: most of models have relatively good performances in Condition 1 and 3, while they seem to have less predicting power in Condition 4 and 5.

Figure 1. Comparing Model Fit across

Exponential, Hyperbolic, Quasi-hyperbolic, Tradeoff, DRIFT, ITCH model

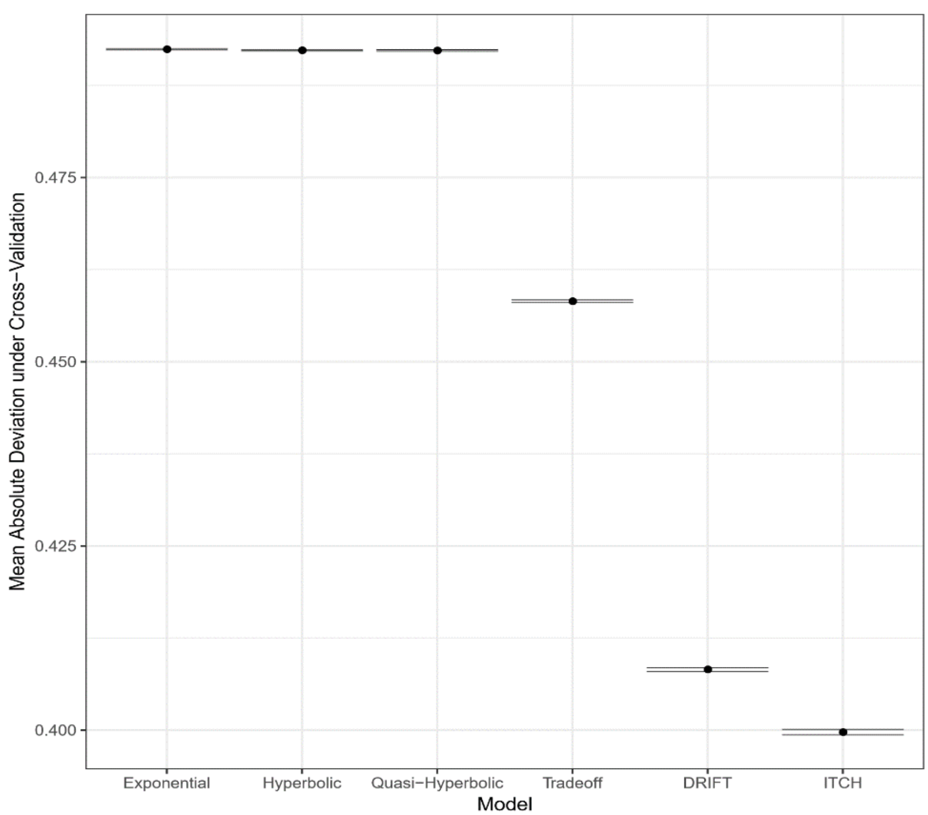
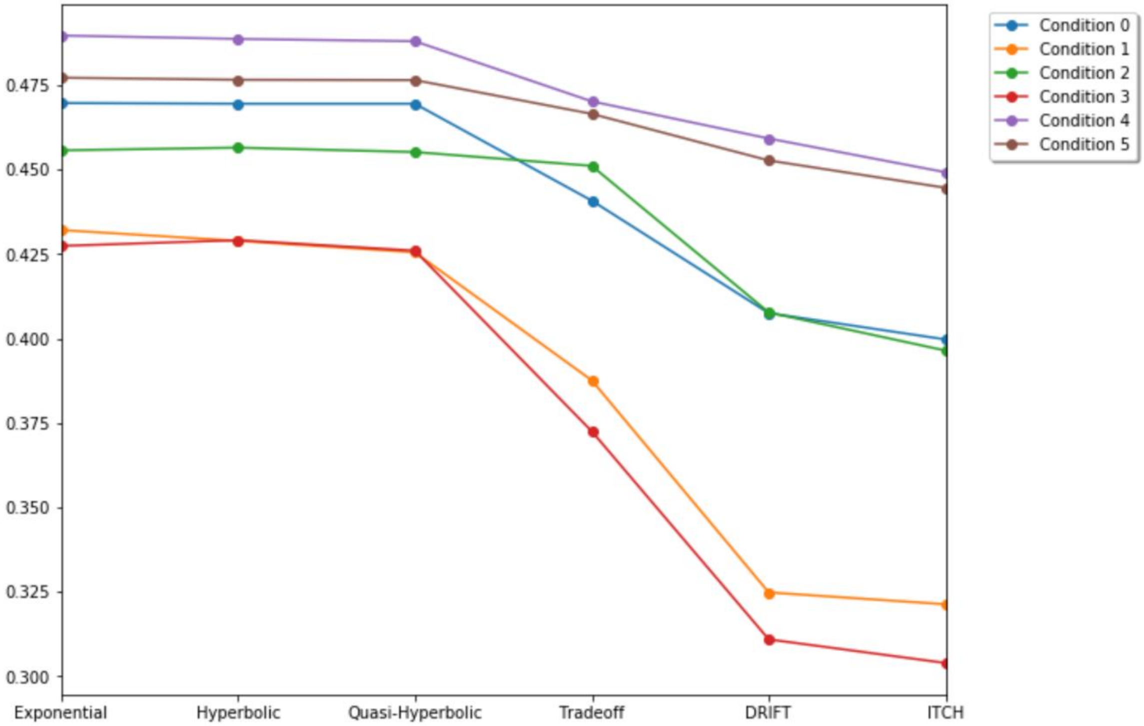


Figure 2. Comparing Model Fit (MAD) across Conditions



Examining the Terms in the ITCH Model

As showed in the above results, the ITCH model has the best performance among all models. The ITCH consists of five terms: the relative time term ($\frac{t_2-t_1}{t^*}$), the absolute time term $(t_2 - t_1)$, the relative money term ($\frac{x_2-x_1}{x^*}$), the absolute money term $(x_2 - x_1)$, and the constant term. Therefore, in order to figure out the reasons why the ITCH model has the best model fit, I examine the characteristics and significance of each term in the ITCH model by backward selection techniques. To be specific, I construct new models by removing one or two similar terms in the ITCH model at a time, and then compare the new model's performance with the original ITCH model. The following are the models without one or two terms:

ITCH Model: *P(LL) = L(*$\beta_I + \beta_{xA}(x_2 - x_1) + \beta_{xR}\frac{x_2-x_1}{x^*} + \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2-t_1}{t^*}$*)*

Removing Relative Time Term: *P(LL) = L(*$\beta_I + \beta_{xA}(x_2 - x_1) + \beta_{xR}\frac{x_2-x_1}{x^*} + \beta_{tA}(t_2 - t_1)$*)*

Removing Absolute Time Term: *P(LL) = L(*$\beta_I + \beta_{xA}(x_2 - x_1) + \beta_{xR}\frac{x_2-x_1}{x^*} + \beta_{tR}\frac{t_2-t_1}{t^*}$*)*

Removing Relative Money Term: *P(LL) = L(*$\beta_I + \beta_{xA}(x_2 - x_1) + \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2-t_1}{t^*}$*)*

Removing Absolute Money Term: *P(LL) = L(*$\beta_I + \beta_{xR}\frac{x_2-x_1}{x^*} + \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2-t_1}{t^*}$*)*

Removing Relative Terms: *P(LL) = L(*$\beta_I + \beta_{xA}(x_2 - x_1) + \beta_{tA}(t_2 - t_1)$*)*

Removing Absolute Terms: *P(LL) = L(*$\beta_I + \beta_{xR}\frac{x_2-x_1}{x^*} + \beta_{tR}\frac{t_2-t_1}{t^*}$*)*

Removing Constant Term: *P(LL) = L(*$\beta_{xA}(x_2 - x_1) + \beta_{xR}\frac{x_2-x_1}{x^*} + \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2-t_1}{t^*}$*)*

Table 2. Manipulating ITCH Model by Removing part(s)

Model Manipulation: Removing ...	Model Fit (Mean Absolute Deviations/MAD)					
	Pooled Data	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5
Baseline / ITCH	0.3997	0.3213	0.3899	0.3158	0.4527	0.4410
Relative Time Term	0.4034	0.3225	0.3924	0.3172	0.4619	0.4460
Absolute Time Term	0.4034	0.3240	0.3948	0.3197	0.4562	0.4440
Relative Money Term	0.4524	0.3963	0.4578	0.3938	0.4761	0.4779
Absolute Money Term	0.4063	0.3320	0.3933	0.3277	0.4572	0.4464
Relative Terms	0.4554	0.3976	0.4592	0.3950	0.4839	0.4829
Absolute Terms	0.4099	0.3350	0.3974	0.3327	0.4603	0.4493
Constant Term	0.4318	0.4097	0.4199	0.4029	0.4585	0.4466

Since all models' performances are fairly stable across conditions, I focus on comparing the model fit resulted from the pooled data, which includes all observations from the five conditions. As Table 2 shows, compared to the original ITCH model, by removing the two relative terms, the MAD increased a lot, which suggests that the relative terms are of great importance; the MAD also increased a lot when removing relative money term, but the model without relative time term doesn't have great changes in model fit, which suggests that the relative money term matters a lot. Besides, when removing the constant term, the MAD increased, which suggests the constant term also plays a significant row in this model.

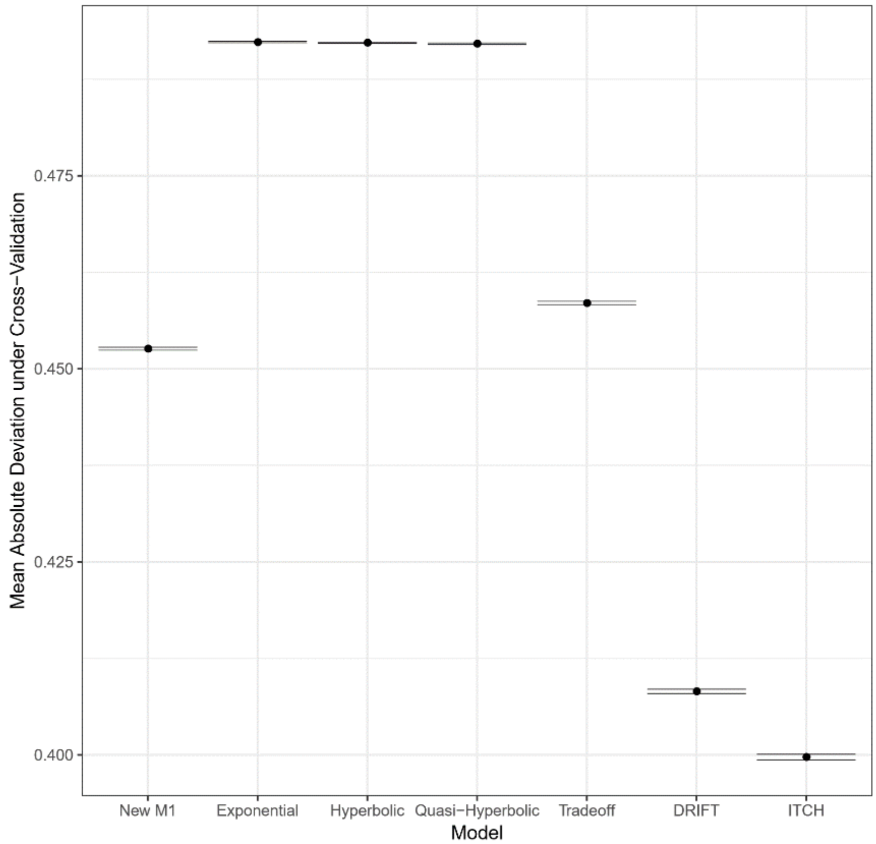
Developing New Models

After examining each term in the ITCH model, I have a better understanding of the potential mechanism of how the ITCH model predicts individuals' intertemporal choice. In order to further examining why delay discounting models don't outperform the ITCH model, I try to develop new models which are nested in the ITCH model and some delay discounting models, such as Exponential model. By comparing the new models with the original six models, I am able to tell what makes the differences.

The analyses above suggest that the relative terms, especially the relative money term, matter a lot. Does a model without any relative term have the worst performance? What if it is compared to the delay discounting models? Therefore, I begin with a new model which has no relative terms but the two absolute terms and a new term: $d_2 - d_1$, which is coded as 0 if T1 > 0 and T2 > 0, or 1 if T1 = 0 and T2 > 0. When it is coded as 0, it means both the two options happen in the future; when it is coded as 1, it means one option happens today while the other happens in the future. The new term tries to play a similar role with the original relative time term. As Figure 3 shows, the new model 1, which has a similar performance as the Tradeoff model, outperforms the three delay discounting models but still underperform the ITCH model.

***New Model 1*:** $P(LL) = L(\beta_I + \beta_2(x_2 - x_1) + \beta_3(t_2 - t_1) + \beta_4(d_2 - d_1))$

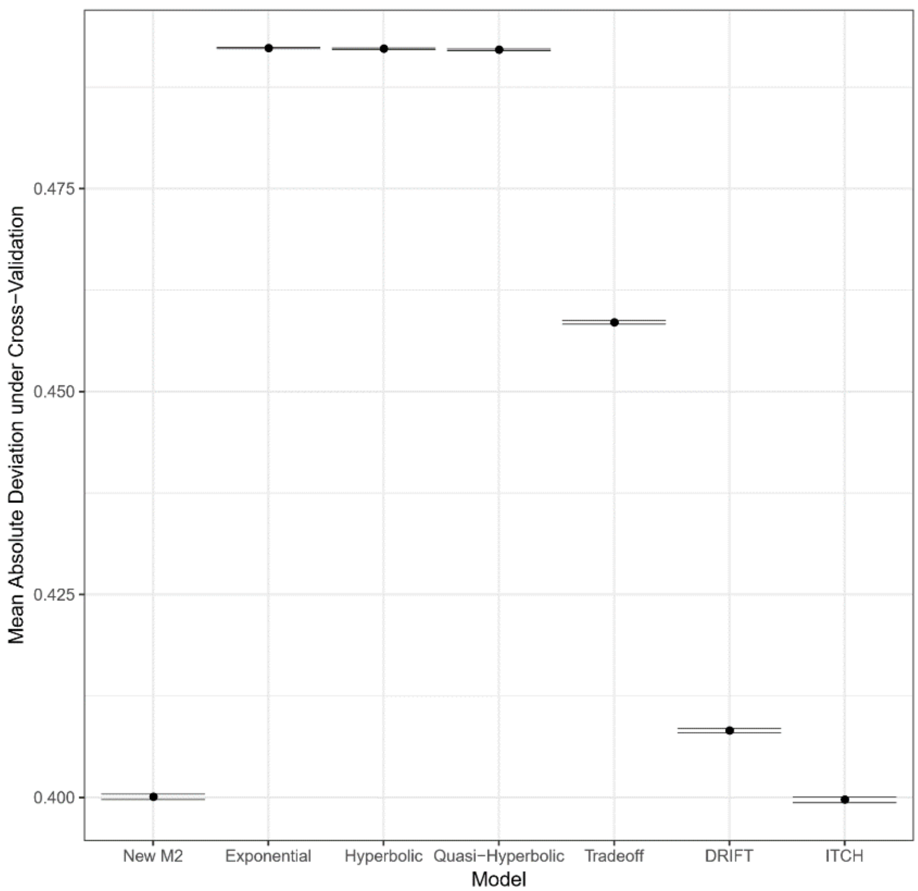
Figure 3. Comparing New Model 1 with Other Six Models



Since the new model 1 doesn't outperform the ITCH model, I try new model 2 which is added the relative money term, which is the most important term in the ITCH model. As Figure 4 shows, it has the exact same performance as the original ITCH model, which suggests that the new time term does play the same role as the old relative time term.

***New Model 2*:** *P(LL)* = $L(\beta_I + \beta_2(x_2 - x_1) + \beta_3 \frac{x_2 - x_1}{x^*} + \beta_4(t_2 - t_1) + \beta_5(d_2 - d_1)$

Figure 4. Comparing New Model 2 with Other Six Models

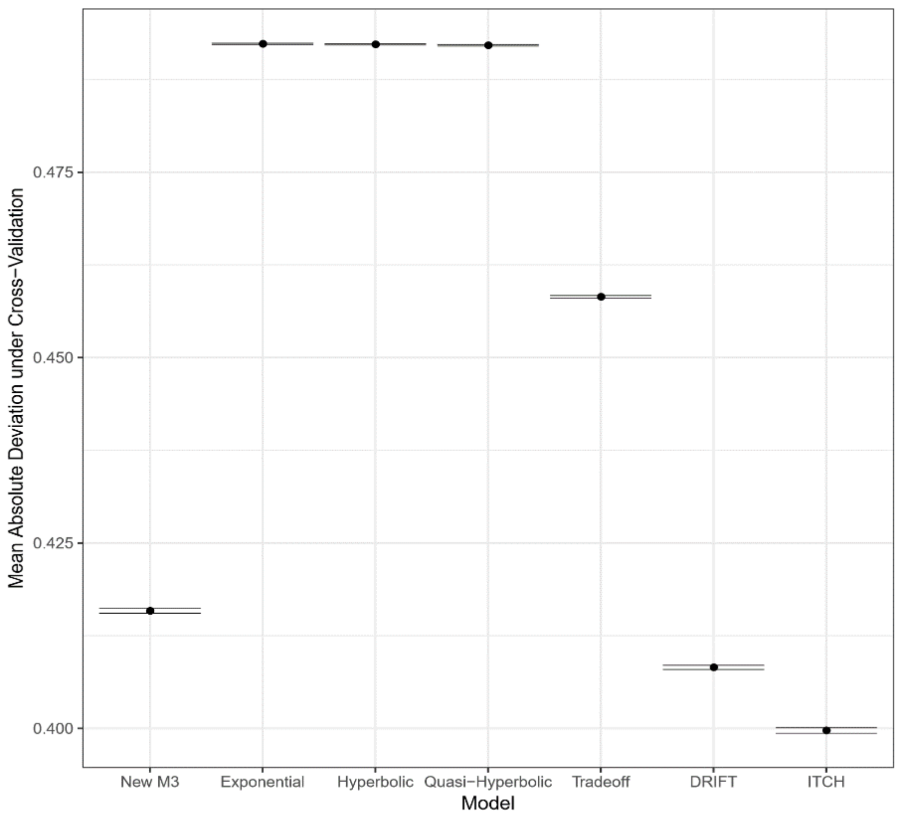


In the new model 3, I add the exponential discounting term to the new model 2. As Figure 5 shows, it has better performance than the delay discounting models, but doesn't outperform the ITCH model. It could be due to overfitting.

***New Model 3*:**

$P(LL) = L(\beta_I\ (x_2\delta^{t2} - x_1\delta^{t1}) + \beta_2(x_2 - x_1) + \beta_3 \frac{x_2 - x_1}{x^*} + \beta_4(t_2 - t_1) + \beta_5(d_2 - d_1)$

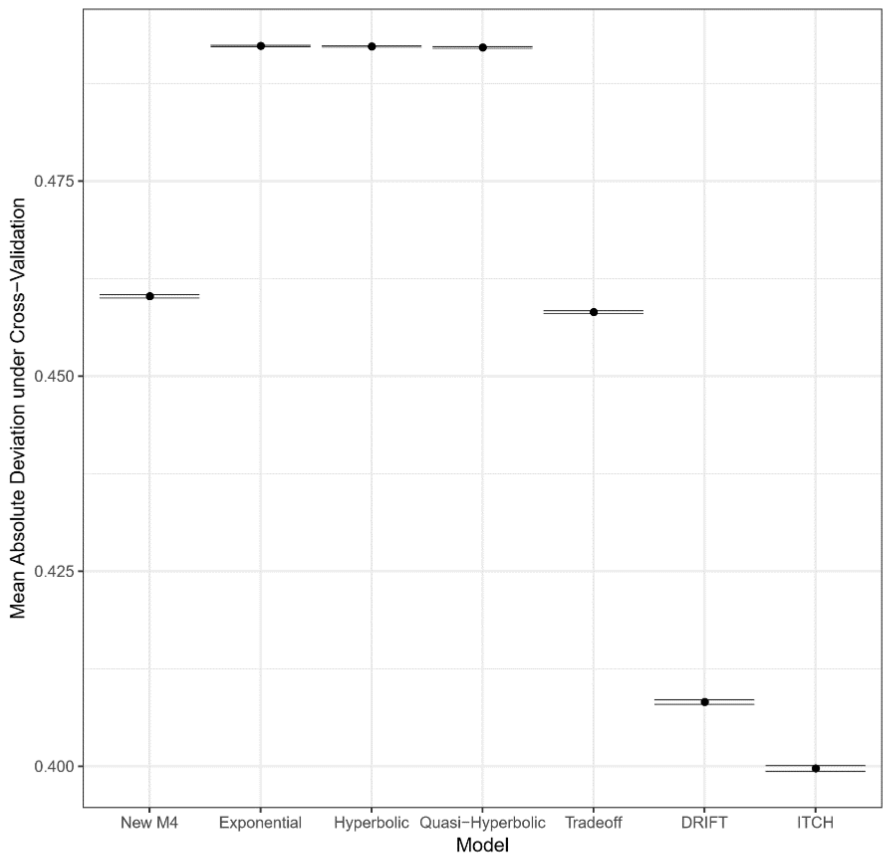
Figure 5. Comparing New Model 3 with Other Six Models



In the new model 4, I only include the exponential discounting term and the new time term. As Figure 6 shows, it has a similar performance as the Tradeoff model, but doesn't outperform the other two heuristic models.

***New Model 4*:** $P(LL) = L(\beta_I\ (x_2{}^{\alpha}\delta^{t2} - x_1{}^{\alpha}\delta^{t1}) + \beta_2(d_2 - d_1))$

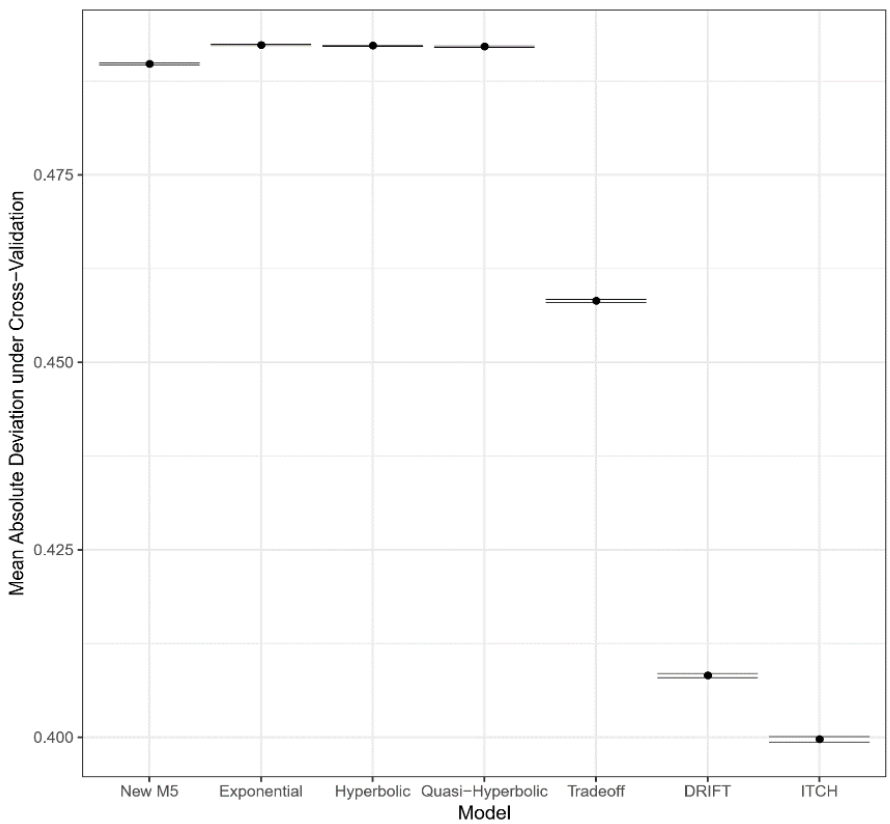
Figure 6. Comparing New Model 4 with Other Six Models



In the new model 5, I include an exponential term which is different from the basic form but has another representation, and the new time term. As Figure 7 shows, it has a similar performance as the delay discounting models, but doesn't outperform the other three heuristic models.

***New Model 5*:** $P(LL) = L(\beta_I\,[(\beta_2 x_2 + \beta_3 \frac{x_2}{x_1})\,\delta^{t2-t1}) - x_1] + \beta_4(d_2 - d_1))$

Figure 7. Comparing New Model 5 with Other Six Models



In the new model 6, I focus on examining the exponential term. I construct an absolute and a relative format for the exponential terms. As Figure 8 shows, it outperforms all three delay discounting models but doesn't outperform the two heuristic models: the ITCH model and the DRIFT model.

***New Model 6*:**

$P(LL) = L(\beta_1(v_2 - v_1) + \beta_2 \frac{v_2 - v_1}{v^*})$ where $v_1 = x_1\delta^{t1}$, $v_2 = x_2\delta^{t2}$, and $v^* = \frac{v_2 + v_1}{2}$.

Figure 8. Comparing New Model 6 with Other Six Models