Intertemporal Choice - Tradeoff between Money and Time

Xi Chen

June 4, 2018

Abstract

Intertemporal choices refer to decisions involving consequences and relative preferences that occur at different time points. Several models have been proposed to address individuals' decision between money and time. These models can be generally categorized into two groups: delay discounting models and heuristic models. This research compares three delay discounting models: exponential model, hyperbolic model, quasi-hyperbolic model, and three heuristic models: ITCH model, Trade-off model, and DRIFT model. This research also examines the characteristics of the ITCH model and tries to develop new models which are nested on the best model(s). Nine new models are proposed based on various focuses and characteristics of intertemporal choice.

Keywords: intertemporal choice, delay discounting models, heuristic models

1 Introduction

Intertemporal choices refer to decisions involving consequences and relative preferences - tradeoff between cost and benefits - that occur at different time points (Frederick, Loewenstein, & O'Donoghue, 2002). For example, a choice with two alternative options: $10 today versus $20 in a week. These intertemporal decisions involving tradeoff between money and time have been extensively studied across several academic domains including economics, psychology, marketing, decision-making science, and recently neuroscience. Intertemporal choices have several alternative names in these domains: *intertemporal discounting*, *time preference, delay discounting,* and *time discounting* (Doyle, 2013).

1.1 Discounted Utility Theory

The most widely used theoretical framework for studying intertemporal choices is the theory of discounted utility (Samuelson, 1937; Frederick et al., 2002), which uses a *discount rate* to address the lose value of delayed consequences. The discounted utility theory states that, the more delayed the consequence is, the more value the rewards lose, i.e., the higher the rewards are discounted. In other words, the higher the discount rate is, the greater the preference for earlier choices over later choices (Chabris, Laibson, & Schuldt, 2010). Discount rates can also be considered as a measure of impatience from a psychological perspective: the higher the discount rate is, the more likely that an individual would choose the earlier choice, which indicates the less patient he or she is.

1.2 Measuring Intertemporal Choices

Past research has measured intertemporal choices by eliciting preferences over various alternative choices obtained at different times. To be specific, the common experimental paradigm is *money-earlier-or-later* (MEL) task, in which individuals choose between a smaller, earlier monetary reward and a larger, later monetary reward (Cohen, et al., 2015). The example mentioned at the beginning - $10 today versus $12 in a week - is a typical MEL task. Based on the MEL task paradigm, two main categories of methods are employed in measuring intertemporal choices: choice based and matching based (Urminsky & Zauberman, 2015).

1.2.1 Choice Based Measure

The choice-based measure (e.g., Kirby & Marakovic, 1996; Tanaka, Camerer, & Nguyen, 2010) presents individuals with a series of MEL tasks to elicit discount rates. For example, $10 today versus $12 in two weeks, $10 today versus $13 in two weeks. By increasing the monetary amount in the following binary choices, researches can identify the point in which individuals switch from the smaller, earlier option to the larger, later option, so that the discount rate can be derived. In this experiment paradigm, both monetary and time delay amount can be incremented to examine the switching point according to various experiment designs.

1.2.2 Matching Based Measure

Instead of presenting a series of MEL tasks, the matching based measure (e.g., Thaler, 1981; Malkoc & Zauberman, 2006) provides individuals with a sooner option such as $10 today, and then ask for their desired monetary amount at a later time such as in two weeks, so that they are indifferent between the sooner and the later option. By asking individuals for a single response of the equivalent amount, researchers can derive the discount rate as well.

2 Modeling Intertemporal Choices

A vast literature in economics, psychology, and business have explained the way individuals trade off time and money by proposing different models with various characteristics. These models can be generally categorized into two groups: delay discounting models and heuristic models. The delay discounting models, which are based on the discounted utility theory, have dominated the research in intertemporal choices for a long time. In this paper, three important and typical delay discounting models are discussed: exponential model, hyperbolic model, and quasi-hyperbolic mode. The heuristic models, which are more recent developed, arouse researchers' interest due to the ability to address intertemporal choice anomalies better than the standard economic models. Three heuristic models are discussed in this paper: ITCH model (Marzilli, Ericson, et al., 2015), Trade-off model (Scholten & Read, 2010), and DRIFT model (Read, Frederick, & Scholten, 2013).

Since each model has various versions in the past literature, this paper uses the versions in Marzilli Ericson et al., (2015). The outcome is a binary variable: 1 represents the larger, later option is chosen, 0 represents the smaller, sooner option is chosen. In each of these models, the notation, *L(z)*, represents the inverse logistic function of z:

$$L(z) = (1 + e^{-z})^{-1}$$

P(LL) represents the probability that individuals prefer the larger, later option over the smaller, sooner option. The variable *a* represents the logistic scaling parameter. The notation *I(x)* represents the indicator variable that is 1 if x is true and 0 otherwise. x_1 and x_2 are the variables of monetary amount; t_1 and t_2 are the variables of time amount.

2.1 Delay Discounting Models

2.1.1 Exponential Model

The first well-known standard economic model is the Discounted Utility Model (Samuelson, 1937), which is also called Exponential discounting model. In this model, intertemporal preferences are governed by the discounted utilities of the options. To be specific, the utilities are exponentially discounting as a function of the delays to the outcomes. Exponential discounting model, which assumes the constant discount rate, ensures consistency in the treatment of time, so that intertemporal choices will not change merely due to time changes (Strotz, 1955). The functional form of the Exponential model is:

$$P(LL) = L(a\ (x_2\delta^{t2} - x_1\delta^{t1}))$$

2.1.2 Hyperbolic Model

Hyperbolic discounting model is motivated by the common difference effect (Loewenstein & Prelec, 1992). Instead of assuming that constant exponential discounting rate as the Exponential model, this model suggests that there is more discounting over an earlier interval that the later interval. The options' values are hyperbolically discounted as a function of the delay to the outcomes. The functional form of the Hyperbolic model is:

$$P(LL) = L(a\ (x_2(1+\alpha t_2)^{-1} - x_1(1+\alpha t_1)^{-1}))$$

2.1.3 Quasi-hyperbolic Model

Quasi-hyperbolic discounting model, which is nested on the Exponential model, employs new features to account for the *present bias* (O'Donoghue & Rabin, 1999): there is more discounting over an interval that begins now than one that begins later. The functional form of the Quasi-hyperbolic model is:

$$P(LL) = L(a\ (x_2 \beta^{I(t2>0)} \delta^{t2} - x_1 \beta^{I(t1>0)} \delta^{t1}\ ))$$

2.2 Heuristic Models

Since standard economics models cannot account for some empirical regularities – intertemporal choices anomalies, researchers develop heuristic models. The anomalies of intertemporal choices (Loewenstein & Prelec, 1992; Laibson, 1997) include but are not limited to: (a) *reversal of preference*: individuals are more impatient for shorter time horizons than for longer intervals; for example, an individual prefers $10 today over $12 tomorrow, but he/she prefers $12 in a year plus a day over $10 in a year (Benhabib, Bisin, & Schotter, 2010); (b) *magnitude effect*: instead of constant discounting rate, the empirical discounting rate is observed to decline with the amount to be discounted (Thaler, 1981).

2.2.1 ITCH Model

ITCH model, which stands for *intertemporal choice heuristic* model, is based on psychological principles rather than economic theory, and inspired by attributed-based models. As Ericson et al., (2015) argues that, "the ITCH model implement four basic psychological principles: (a) each option is compared to a reference point (Kahneman & Tversky, 1979); (b) comparisons are performed in both absolute terms (by subtraction) and relative terms (by division; Thurstone, 1927); (c) comparisons are performed independently along the monetary and time dimensions (Lichtenstein & Slovic, 1971); and (d) the results of these comparisons are then aggregated linearly using a set of decision weights (Busemeyer & Townsend, 1993)." The functional form of the ITCH model is:

$$P(LL) = L(\beta_I + \beta_{xA}(x_2 - x_1) + \beta_{xR}\frac{x_2 - x_1}{x^*} + \beta_{tA}(t_2 - t_1) + \beta_{tR}\frac{t_2 - t_1}{t^*}), \text{ where } x^* = \frac{x_2 + x_1}{2}.$$

2.2.2 Trade-off Model

Scholten & Read (2010) develop the Trade-off model in which "people make intertemporal choices by weighing how much more they will receive or pay if they wait longer against how much longer the wait will be, or, conversely, how much less they will receive or pay if they do not wait longer against how much shorter the wait will be." The Trade-off model accounts for several anomalies that the delay discounting models cannot address. The functional form of the Trade-off model is:

$$P(LL) = L(a\,((\frac{\log(1+\gamma_x x_2)}{\gamma_x} - \frac{\log(1+\gamma_x x_1)}{\gamma_x} - k(\log(1+\gamma_x t_2) - \frac{\log(1+\gamma_x t_1)}{\gamma_x})))$$

2.2.3 DRIFT Model

People would like to receive good outcomes immediately rather than wait, and they would need to be compensated for waiting. In order to examine what influences individuals' decision about how much compensation is required for a given wait, Read, Frederick, & Scholten (2013) propose the DRIFT model, which stands for *difference-ratio-interest-finance-time* and its functional form is:

$$P(LL) = L(\beta_0 + \beta_1(x_2 - x_1) + \beta_2\frac{x_2 - x_1}{x^*} + \beta_3 \frac{x_2}{x_1}^{\frac{1}{t2-t1}} + \beta_4(t_2 - t_1))$$

3 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. Each participant was randomly assigned to one of the five conditions described in the following. Each condition has different MEL tasks which varied in framing contexts for the money terms (absolute or relative) and 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)

After being assigned to one of the conditions, each participant was asked to answer 25 MEL questions, and each MEL task has various 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.

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

Table 1 presents descriptive statistics for the five key variables: $Money_{SS}$ refers to the monetary value associated with the smaller, sooner option; $Time_{SS}$ refers to the time amount associated with the smaller, sooner option; $Money_{LL}$ refers to the monetary value associated with the larger, later option; $Time_{LL}$ 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

	$Money_{SS}$	$Time_{SS}$	$Money_{LL}$	$Time_{LL}$	$LaterOptionChosen$
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 order to compare the performance of each model, the measure being used 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.

In the following presentation of each model's functional form, $Money_{SS}$ is denoted as x_1, $Time_{SS}$ is denoted as t_1, $Money_{LL}$ is denoted as x_2, and $Time_{LL}$ is denoted as t_2.

4 Analysis and Results

4.1 Comparing Model Fit Across Six Models

To begin with, a comparison of model fit is conducted across six 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

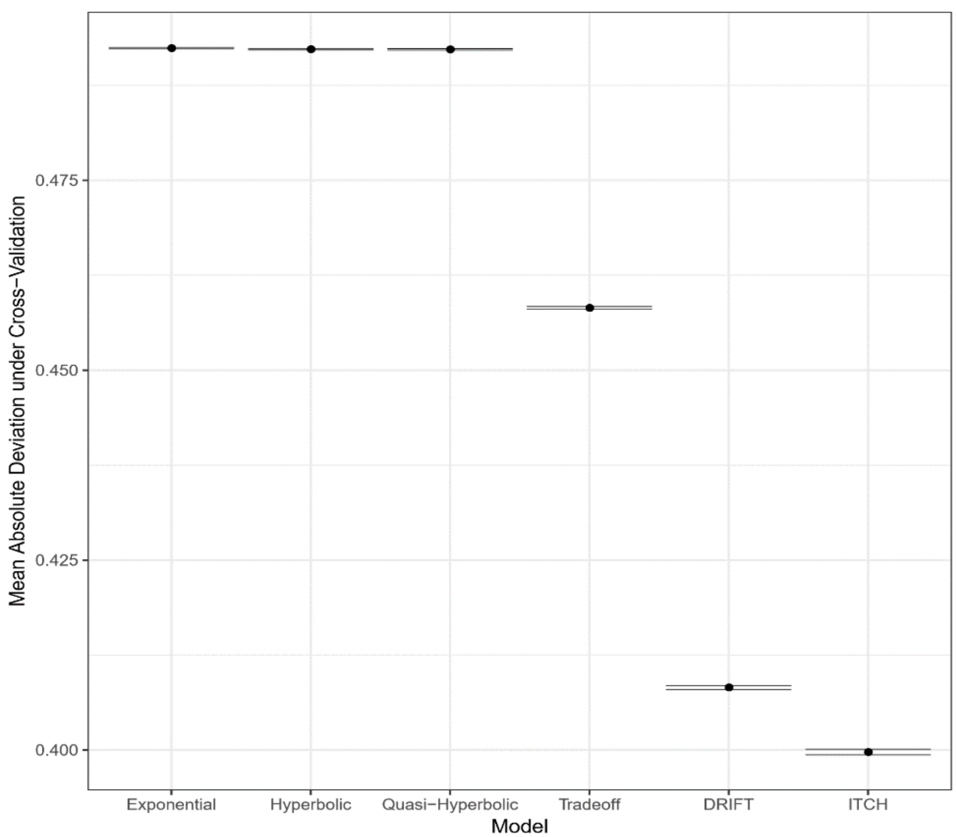
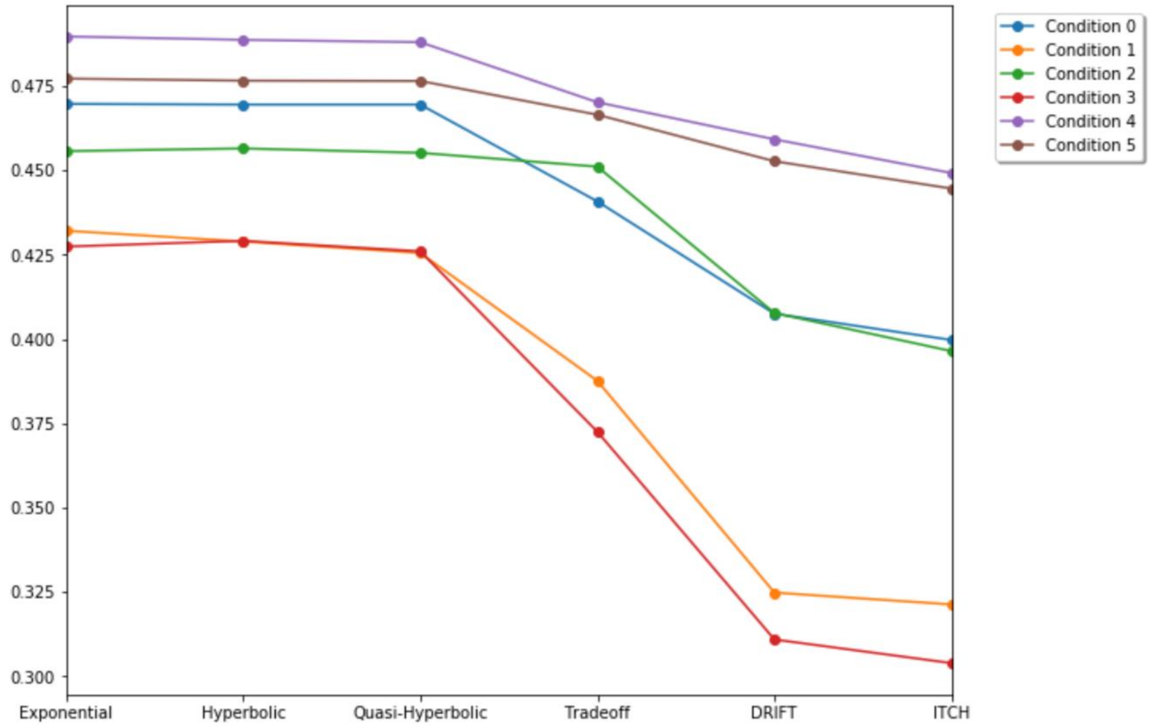


Figure 2. Comparing Model Fit (MAD) across Conditions



4.2 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 to the original ITCH model. The following shows the actual models that I examine:

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 quite stable across conditions, I focus on comparing the model fit generated 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.

4.3 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.

In addition, the ITCH model doesn't seem to have an explicit term to address the issue of present bias, which means that people have the tendency to prefer the options that are closer to the present time when considering options between two future moments (O'Donoghue, & Rabin, 1999). Therefore, I try a new time term: $d_2 - d_1$, which is coded as 0 if t_1> 0 and t_2 > 0, or 1 if t_1 = 0 and t_2 > 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.

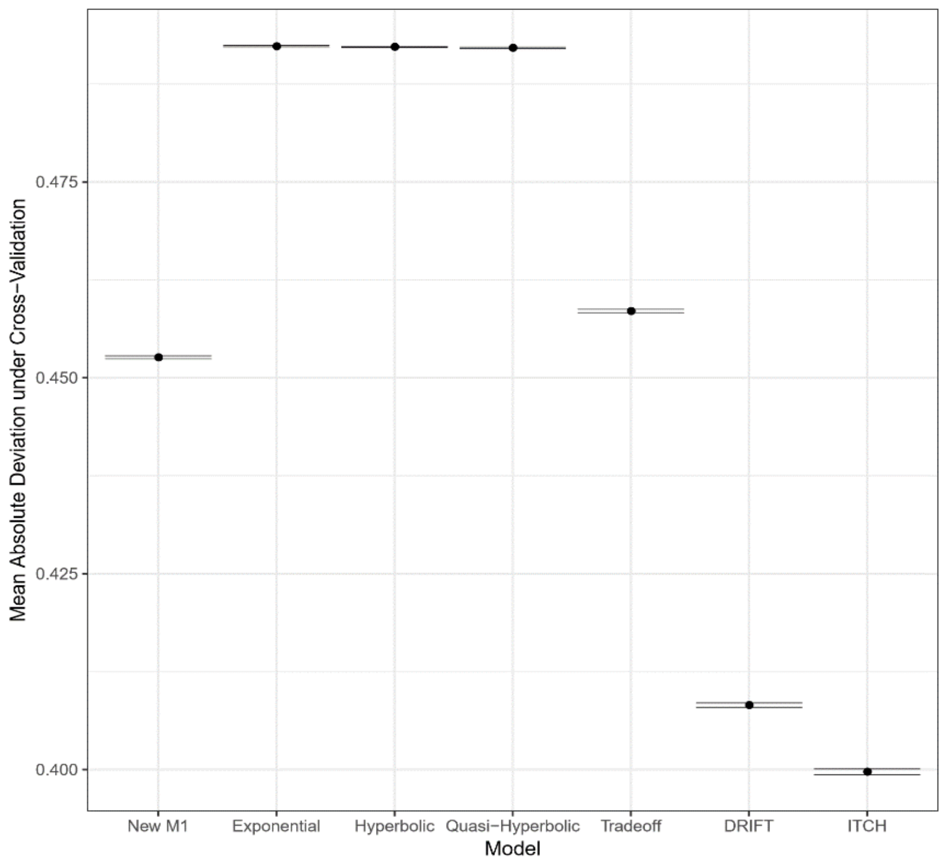
From the results of examining ITCH model's terms by backward selection, I find that the absolute time term doesn't have much power. Therefore, in the following new models that I construct, I replace the absolute time term $t_2 - t_1$ with the new time term $d_2 - d_1$.

The following presents nine new models and their performances relative to the original six models.

The above analyses suggest that the relative terms are important. Does a model without any relative term have the worst performance? The worst performance may serve as a baseline for the model fit. Therefore, I begin with a new model which has no relative terms but the two absolute terms and the new time term.

***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



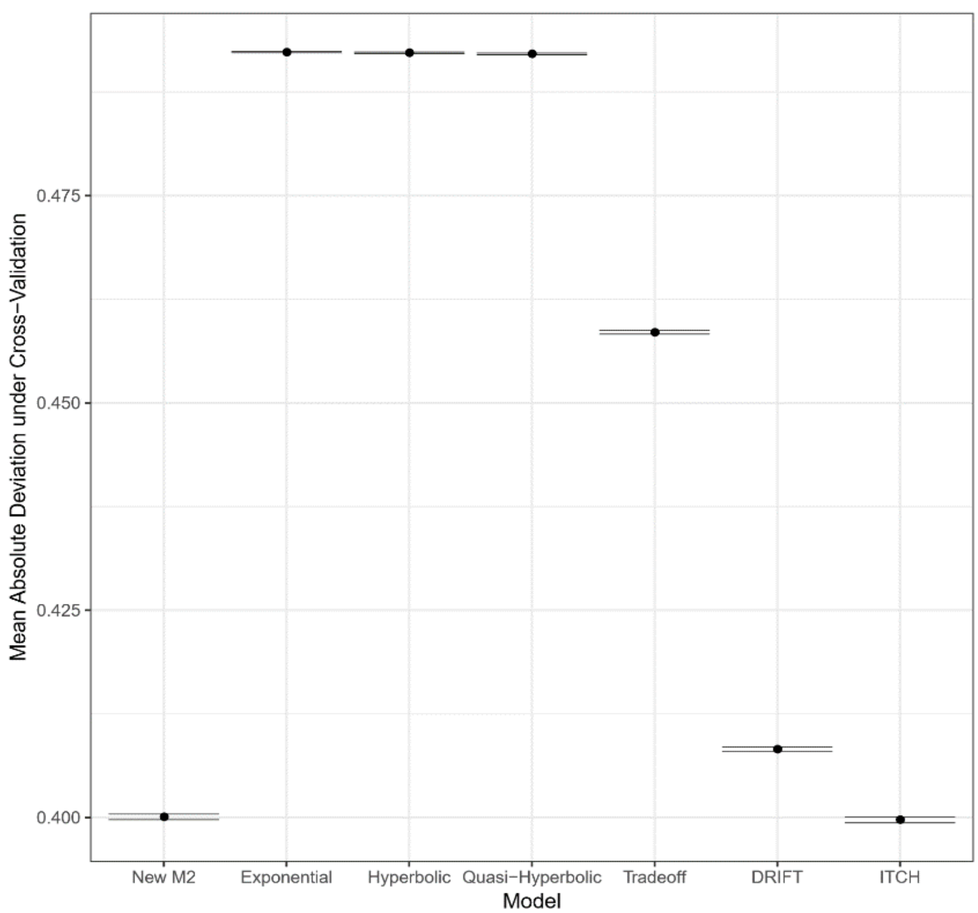
As the following Figure[2] 3 shows, the new model 1, which has a similar performance as the Tradeoff model, outperforms the three delay discounting models but still underperform the ITCH model. The result suggests that one or both of the relative terms could be important.

[2] In all the following figures, the error bars are represented by two horizontal lines for each MAD results.

Since the result from the new model 1 suggests that at least one of the relative terms matters, I add the relative money term into the new model 2. It is because the results from examining each term in the ITCH model by backward selection show that the relative money term is the most important term. As Figure 4 shows, the new model 2 has a very close model fit when compared to the ITCH model. This suggests that the new time term is working, and it plays a role as good as the old relative time term. Therefore, in the following new models, I keep the new time term $d_2 - d_1$.

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

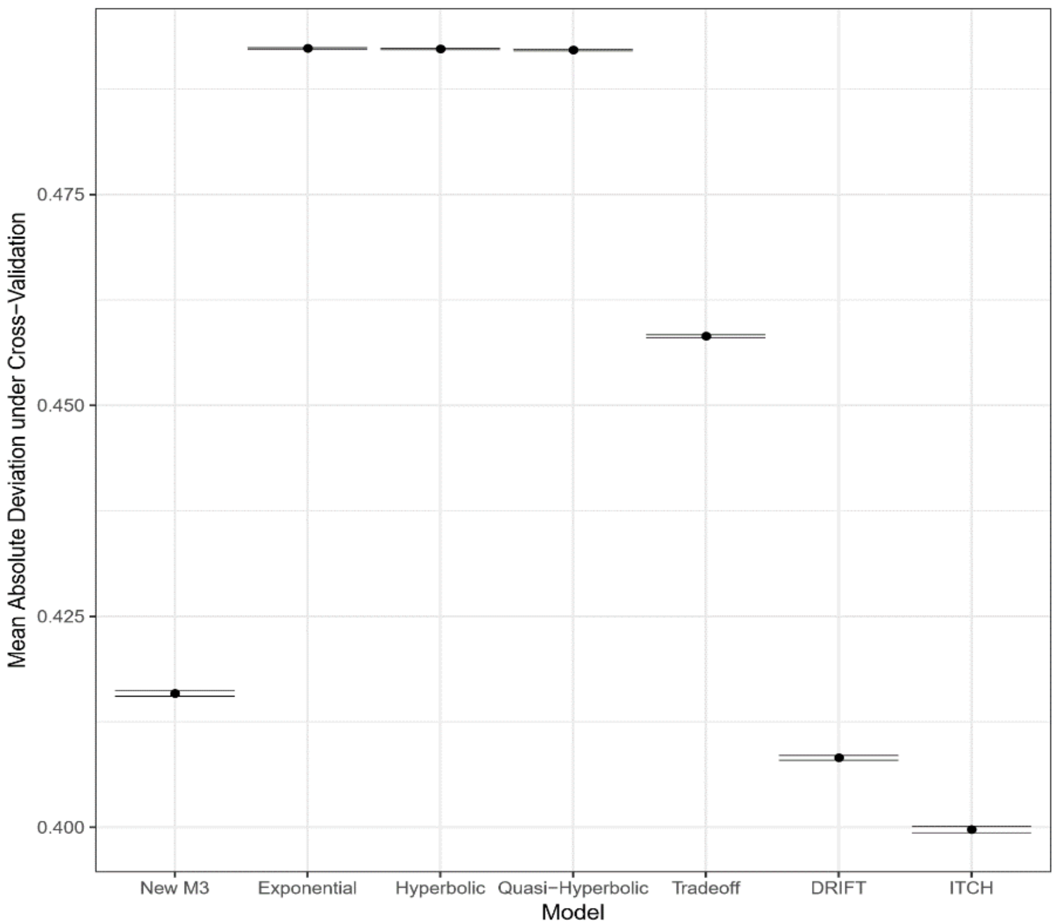


For the new model 3, I add the exponential discounting term to the basis of the new model 2. As Figure 5 shows, it has better performance than the delay discounting models, but doesn't outperform the ITCH model. The worsen performance may be due to overfitting since there are five terms.

***New Model 3*:**

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

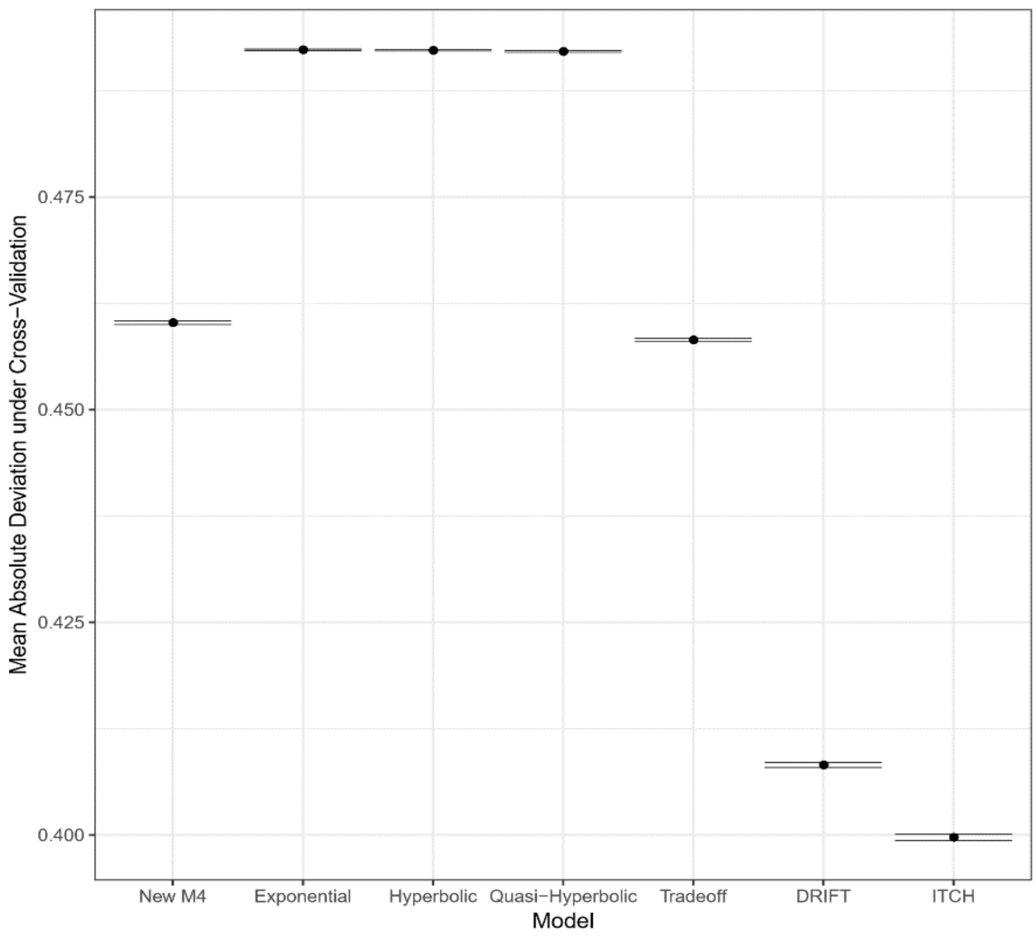
Figure 5. Comparing New Model 3 with Other Six Models



For the new model 4, I only include the exponential discounting term and the new time term, because I want to address the potential overfitting problem in the new model 3. As Figure 6 shows, it has a similar performance as the Tradeoff model, but doesn't outperform the other two heuristic models. This suggests that only the exponential term and the present bias term may not be able to capture all the characteristics of individuals' intertemporal choice.

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

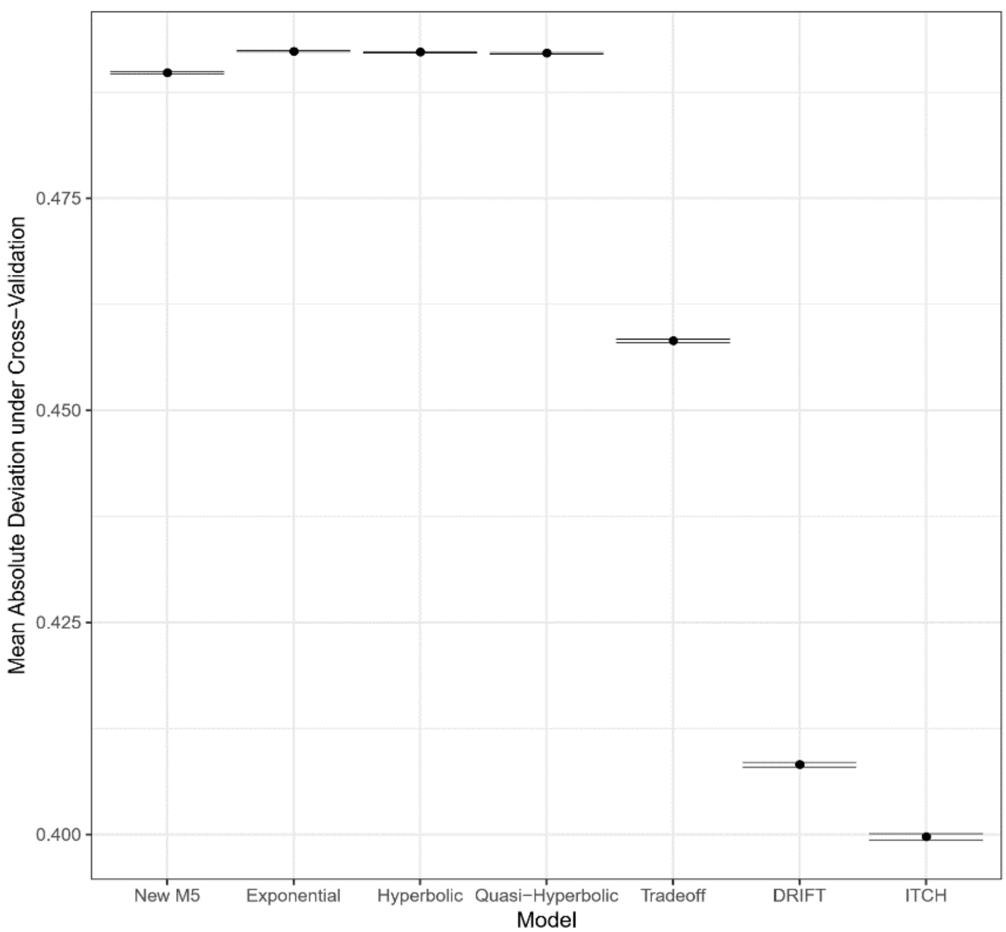
Figure 6. Comparing New Model 4 with Other Six Models



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. It has the worst performance among the new models so far, which suggests that the original form of exponential term is better than this form.

***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

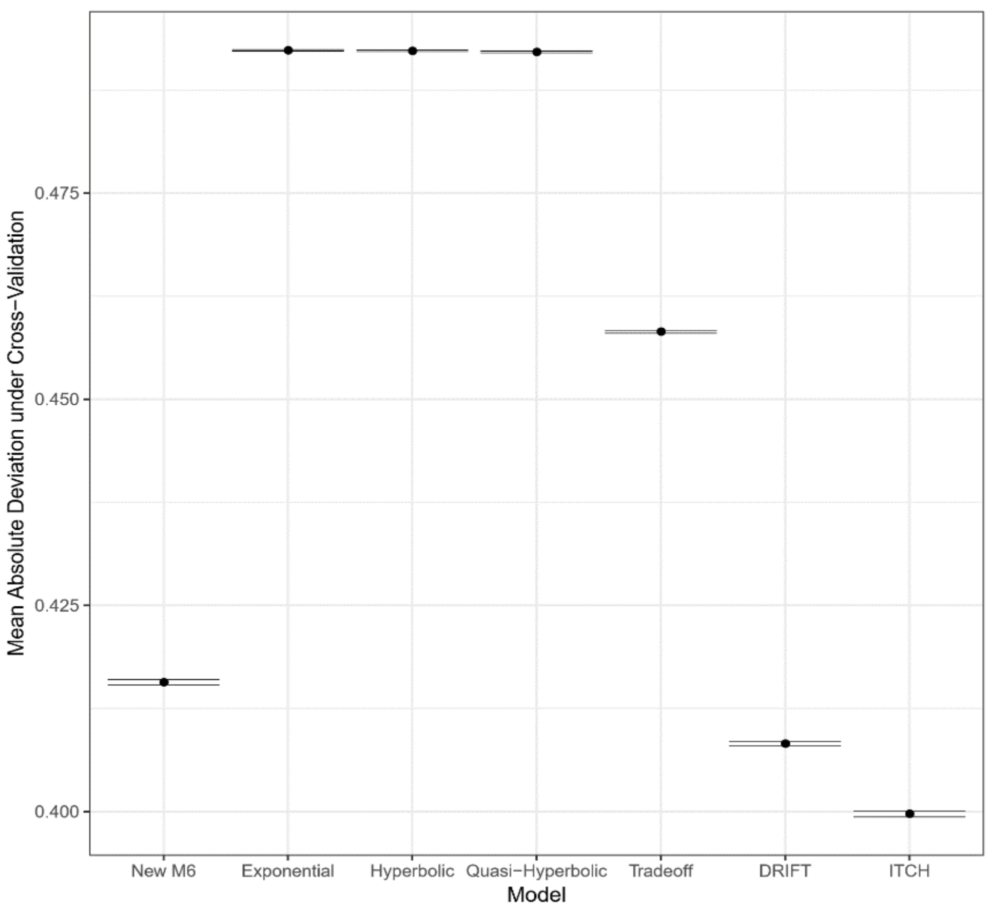


In the new model 6, I continue examining the variants of the exponential term. Specifically, I construct an absolute form and a relative form for the exponential terms. As Figure 8 shows, it outperforms all three delay discounting models but doesn't outperform the two heuristic models: the ITCH model and the DRIFT model. However, it is much better that the new model 5, which is a good sign. This model has several advantages: firstly, it only has two terms; secondly, it is nested on the exponential 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 $v^* = \frac{v_2 + v_1}{2}$.

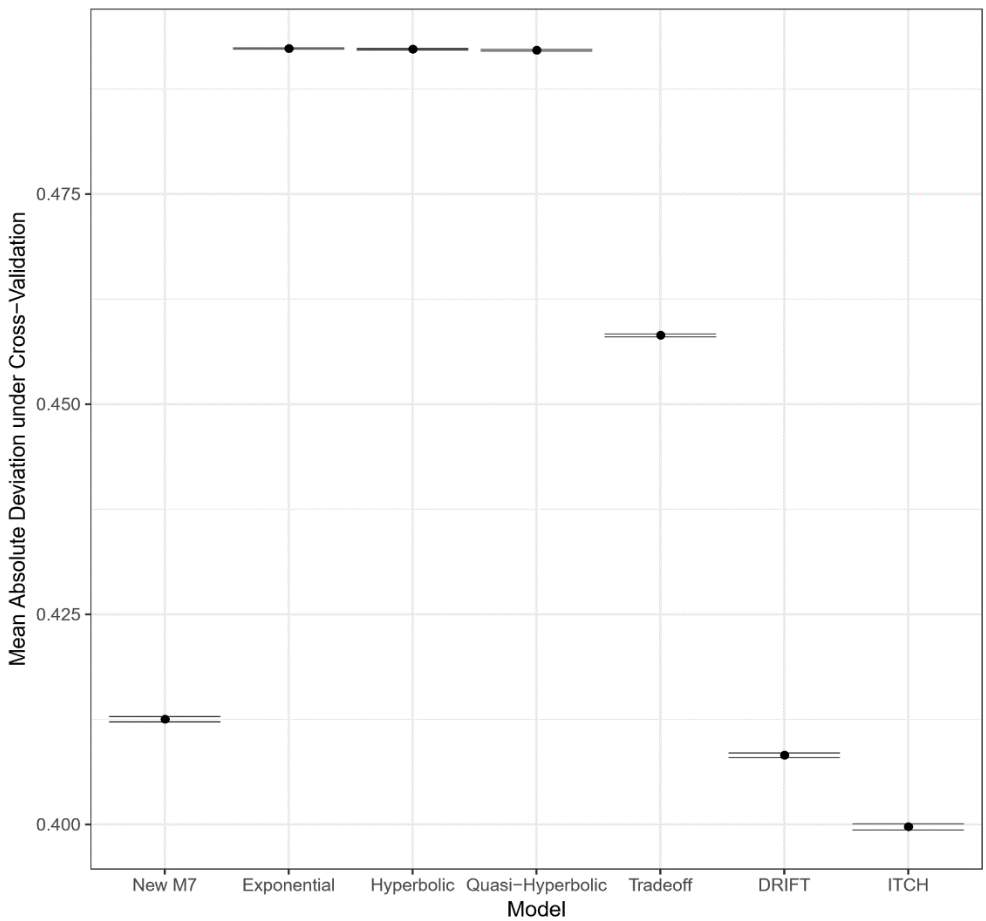
Figure 8. Comparing New Model 6 with Other Six Models



For the new model 7, I add the new time term into the new model 6, and I want to see if the time term is helpful in this case. As Figure 9 shows, it has a slightly better performance compared to the new model 6. However, it doesn't outperform the ITCH model and the DRIFT model.

New Model 7: $P(LL) = L(\beta_1(v_2 - v_1) + \beta_2 \frac{v_2 - v_1}{v^*} + \beta_3(d_2 - d_1))$

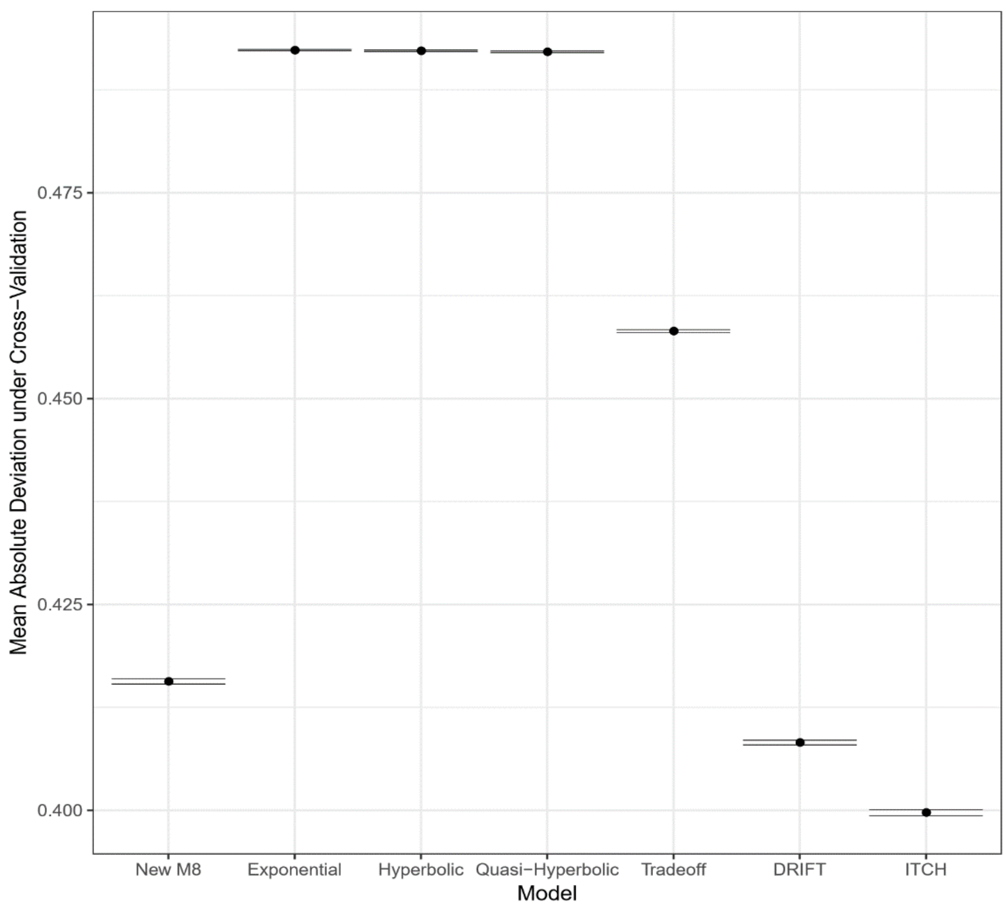
Figure 9. Comparing New Model 7 with Other Six Models



In the new model 8, I only include the relative form of the exponential terms. As Figure 10 shows, it has similar model fit as the new model 6, and it performs a little bit worse than the new model 7. The new model 8 still doesn't outperform the ITCH model and the DRIFT model. However, the result of this model suggests that the relative form of the exponential term is more important than the absolute form of the exponential term. In addition, when comparing this model's output to that in the new model 6, the absolute form of the exponential term seems to be powerless.

***New Model 8*:** $P(LL) = L(\, \beta_1 \frac{v_2 - v_1}{v^*}\,)$

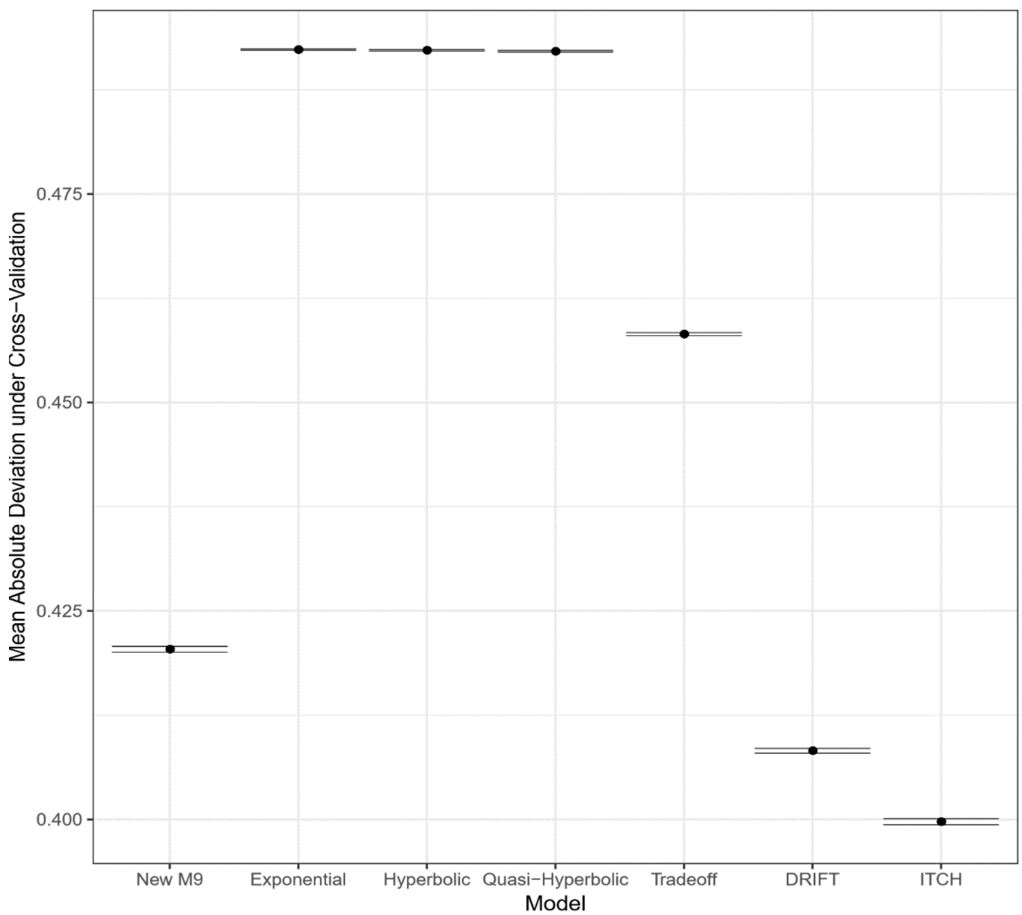
Figure 10. Comparing New Model 8 with Other Six Models



In the new model 9, which is the last new model that I construct, the relative form of the exponential term and the new time term are included. Surprisingly, as Figure 11 shows, it doesn't outperform the new model 8. Compared to the original six models, the new model 9 outperforms all three delay discounting models and one heuristic model, but still doesn't outperform the other two heuristic models: the ITCH model and the DRIFT model.

***New Model 9*:** $P(LL) = L(\beta_1 \frac{v_2 - v_1}{v^*} + \beta_2(d_2 - d_1))$

Figure 11. Comparing New Model 9 with Other Six Models



5 Conclusion

In this paper, I study intertemporal choice, and I examine several classical models which are proposed to address individuals' decisions between money and time. These models can be categorized into two groups: delay discounting models and heuristic models. Specially, I compare three delay discounting models: exponential model, hyperbolic model, quasi-hyperbolic model, and three heuristic models: ITCH model, Trade-off model, and DRIFT model. By virtue of iterative cross-validation analyses, I find that the three heuristic models outperform the three delay discounting models, and the ITCH model has the best performance, which are consistent with the results in Ericson et al., (2015). In order to figure out why the ITCH model has the best performance, I use backward selection techniques to examine the importance of each term at the ITCH model. The results show that the relative terms are important, and the relative money term is especially meaningful.

After having a better understanding of each model, I develop nine new models which focus on capturing different characteristics of intertemporal choice. I firstly try replacing the absolute time term with a term which may address the present bias. The new model 2 has nearly the same performance as the ITCH model, which may suggest that individuals have present bias in their intertemporal choice. I also try to construct several different forms and/or combinations of exponential terms. However, some models with the complicated forms of exponential terms may have the problem of overfitting. Among all models, the result for the new model 7 is especially promising because this model not only take present bias into considerations but also it is nested on the exponential model.

Even though most of the new models outperform three delay discounting models and one heretics model, none of them defeats the ITCH model. The future direction is to continue developing new models which hopefully can have better performance than the ITCH model. In addition, new experiments may be designed to compare models' performance across conditions. Furthermore, since the data used in this paper is generated from the choice-based measure, new experiments may be designed to employ matching-based measure to examine the manner in which individuals discount future monetary options.

References

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior, 69*(2), 205-223.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review*, *100*(3), 432.

Chabris, C. F., Laibson, D. I., & Schuldt, J. P. (2010). Intertemporal choice. In *Behavioural and Experimental Economics* (pp. 168-177). Palgrave Macmillan, London.

Doyle, J. R. (2013). Survey of time preference, delay discounting models.

Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of economic literature*, *40*(2), 351-401.

Kirby, K. N., & Maraković, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic bulletin & review*, *3*(1), 100-104.

Tversky, A. & Kahneman, D. (1979) Prospect Theory: An Analysis of Decision under Risk. In: Econometrica. RePEc:ecm:emetrp:v:47:y:1979:i:2:p:263-91.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, *112*(2), 443-478.

Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, *107*(2), 573-597.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of experimental psychology*, *89*(1), 46.

Marzilli Ericson, K. M., White, J. M., Laibson, D., & Cohen, J. D. (2015). Money earlier or later? Simple heuristics explain intertemporal choices better than delay discounting does. *Psychological science*, *26*(6), 826-833.

Malkoc, S. A., & Zauberman, G. (2006). Deferring versus expediting consumption: The effect of outcome concreteness on sensitivity to time horizon. *Journal of Marketing Research*, *43*(4), 618-627.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, *89*(1), 103-124.

Read, D., Frederick, S., & Scholten, M. (2013). DRIFT: An analysis of outcome framing in intertemporal choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 573.

Samuelson, P. A. (1937). A note on measurement of utility. *The review of economic studies*, *4*(2), 155-161.

Scholten, M., & Read, D. (2010). The psychology of intertemporal tradeoffs. *Psychological review*, *117*(3), 925.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, *23*(3), 165-180.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics letters*, *8*(3), 201-207.

Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and time preferences: linking experimental and household survey data from Vietnam. *American Economic Review*, *100*(1), 557-71.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological review*, *34*(4), 273.

Urminsky, O., & Zauberman, G. (2015). The psychology of intertemporal preferences. *The Wiley Blackwell handbook of judgment and decision making*, 141-181.