An Empirical Specification of Rational Inattention Multinomial Logit (RI- MNL) Model for Panel Datasets

Saeed Shakib (✉ saeed.shakib@mail.utoronto.ca)  
University of Toronto

Khandker Nurul Habib  
University of Toronto

Research Article

Keywords: Rational inattention, Residential location preferences, Discrete choice models, Panel choice datasets, COVID-19 pandemic

Posted Date: August 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1925995/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
An Empirical Specification of Rational Inattention Multinomial Logit (RI-MNL) Model for Panel Datasets

Saeed Shakib M.Sc. (Corresponding Author)
Ph.D. Candidate, Department of Civil & Mineral Engineering
University of Toronto
35 St. George Street, Toronto, Ontario, M4S 1A4, Canada
Email: saeed.shakib@mail.utoronto.ca

Khandker Nurul Habib, Ph.D. P.Eng.
Percy Edward Hart Professor in Civil & Mineral Engineering,
University of Toronto
35 St. George Street, Toronto, Ontario, M4S 1A4, Canada
Email: khandker.nurulhabib@utoronto.ca

FUNDING DECLARATION
This study was funded through an NSERC discovery grant and the Percy Edward Hart professorship fund.

CONFLICT OF INTEREST
All authors declare that they have no conflicts of interest.

AUTHOR CONTRIBUTION STATEMENT
The authors confirm contribution to the paper as follows: study conception and design: K. M. N. Habib; data collection: S. Shakib; model specification: S. Shakib; analysis and interpretation of results: S. Shakib, and K. M. N. Habib, draft manuscript preparation: S. Shakib, and K. M. N. Habib. All authors reviewed the results and approved the final version of the manuscript.
ABSTRACT
Classical discrete choice models are based on Random Utility Maximization (RUM), which assumes that decision makers have complete information on all of the attributes of their choices. Nowadays, this assumption is becoming less relevant with the expansion of the internet and social media, which provide a universal platform that grants anyone access to an endless amount of information. As a result, decision-making environments are transforming into settings in which individuals can be overwhelmed with information sources competing for their attention, causing them to become less attentive to the alternatives in their choice set. In such cases, individuals find themselves in scenarios where they wish to maximize utility but are uncertain about the payoffs associated with each action. There have been recent theoretical advancements in the study of the behaviour of inattentive decision makers in discrete choice contexts with the introduction of the Rational Inattention Multinomial Logit (RI-MNL) model. This paper proposes a closed-form empirical specification in the case of panel choice datasets to supplement the theoretical foundation of the (RI-MNL) model. Following this proposal is a sample application of the proposed specification using data from a panel survey on residential location preferences for the Greater Toronto Area in 2020 and 2021. This paper illustrates how the findings of the empirical model can be interpreted in the context of rational inattention theory.

KEYWORDS
Rational inattention, Residential location preferences, Discrete choice models, Panel choice datasets, COVID-19 pandemic.
1. Introduction

Discrete choice models based on Random Utility Maximization (RUM) theory have received much interest in travel demand modelling in past decades. Classical discrete choice models are based on the multinomial logit model (McFadden, 1973), which assumes that decision makers are rational (i.e., they follow a logical process to make their choice) and fully aware of their choice set, which contains Independent and Irrelevant Alternatives (IIA). Since then, the evolution of discrete choice models has centred on two aspects: 1) Relaxing the IIA assumption to allow similarity between alternatives (e.g. Nested Logit models (Wen and Koppelman, 2001), Generalized Extreme Value models (McFadden, 1978), etc.) 2) Considering consumer taste heterogeneity toward the alternatives in their choice sets (Mixed Logit Models (McFadden and Train, 2000)). In pursuing this research direction for discrete choice models, the assumption that respondents have perfect knowledge of the alternatives in their choice set was accordingly disregarded. However, there have been some recent attempts to drop the assumption of perfect knowledge by incorporating the Rational Inattention (RI) theory (Sims, 2003) into the framework of discrete choice models (Fosgerau et al., 2020; Matêjka and McKay, 2015). RI discrete choice models differ from traditional discrete choice models in that they assume decision makers to be Bayesian agents with prior beliefs on the alternatives in their choice set and exhibit heterogeneous processing of incoming information signals. In the literature, the application of RI discrete choice models has been restricted to using simulated data for validation purposes. This paper presents an empirical specification for Rational Inattention Multinomial Logit (RI-MNL) on panel datasets.

The empirical investigation of the rational inattention discrete choice model in this study is primarily motivated by the rising accessibility of the internet over the past decade. The internet offers its users on-demand access to information, and this information consumes their attention. Widespread access to information leads to a gradual reduction and famine in attention (Simon, 1984). Consequently, it is not unreasonable to expect that this variation in the decision-making environment could lead to alterations in decision-making behaviour. Traditional discrete choice models were developed when decision-makers acquired information on the attributes of the handful of available alternatives via experience. In such circumstances, the assumption that decision-makers are aware of all their alternatives was not utterly irrelevant. Currently, the internet is at the disposal of decision-makers. It allows them to choose which alternatives to focus on and how much information to gather about them. For instance, the rapid growth of smartphone users not only has broadened urban passenger travel mode choices by including new alternatives such as ride-hailing with dynamic prices and shared mobility services but also provides real-time information on network traffic congestion and the estimated arrival times and occupancy level of the transit services. Or, in the context of residential location choice, households that traditionally obtained their choice set through personal experience or word-of-mouth from friends and family are seconds away from accessing online property listings to monitor available options and online forums to obtain detailed information on each favoured neighbourhood. Regardless of the context, these examples share a common aspect in which decision-making settings are expanding in width and depth, creating situations wherein decision makers are unable to absorb and process all available information.

The rational inattention theory describes decision-making behaviour in an information-rich environment when user attention is drained. According to the rational inattention theory,
decision makers not only attempt to maximize their "expected payoffs" but also reduce the cognitive load associated with digesting new information. The term "expected payoff" is adopted here because decision-makers cannot gain the perfect information on the payoffs of each alternative in their choice set. Instead, decision-makers are regarded as Bayesian agents who update their knowledge of the "expected payoff" of the alternatives depending on the information signals they choose to receive and process. Regarding decision makers as Bayesian agents who update their prior beliefs on the alternatives' payoffs in their choice set over time makes rational inattention a suitable approach in situations where panel datasets containing the temporal effect on choice behaviour are available.

This paper proposes a model specification that could be applied to estimate the RI-MNL model in a Markov process for each time frame of panel datasets. The described model is then used on a panel choice experiment dataset gathered in July 2020 and July 2021 to identify the effects of the COVID-19 pandemic on residential location choice behaviour in the Greater Toronto Area (GTA). Residential location choice is generally associated with a wide choice set and many alternative attributes, making it an attention-consuming decision environment, consequently, an appropriate context for applying rational inattention theory.

The next section of this study explores approaches for analyzing panel datasets in the literature. The RI-MNL theoretical model is introduced in the third section. The RI-MNL empirical specification for estimating the panel dataset is described in the fourth section. The findings of the applied RI-MNL panel model on the chosen dataset are discussed in the fifth section. The final section contains closing remarks on this method and suggestions for future direction.

2. Approaches in the Analysis of Panel Choice Datasets

This section explains where the RI-MNL stands regarding current strategies in analyzing panel choice datasets. A brief literature review reveals that panel choice datasets have been analyzed using various methods. This section focuses on the distinctions between these methodologies to highlight the contribution of the RI-MNL panel model. Therefore, we have categorized the analysis of panel choice data into three main approaches.

The first and simplest approach is to conduct parallel analyses of the panel choice dataset for multiple periods and compare the results. Parallel analysis can range from descriptive analysis to discrete outcome and discrete choice models applied to each cycle of panel choice datasets. Jensen et al. (2014) performed a parallel descriptive analysis of the two cycles of choice experiment to capture respondents' perceptions of electric vehicles before and after ownership. Using parallel logistic regression analysis on three cycles of panel data, Meeran et al. (2017) identified changes in the attributes of some selected products. Ambos et al. (2020) employed two distinct MNL models to capture four-year changes in managers' attitudes about internationalization decisions. Hanly and Dargay (2000) merged panel datasets on car ownership decisions and estimated the elasticities for different years to determine the changes in panel data cycles towards various attributes. Parallel analysis is the method for panel data analysis with the least computational complexity. Nevertheless, these methods disregard the possibility of choice persistence in panel datasets by ignoring the existence of state dependence over cycles of the dataset.
The second approach in panel choice analysis uses mixed logit or probit models to estimate discrete choice models assuming that random panel effects are in panel choices. This method is advantageous since it accounts for taste heterogeneity. However, because these models lack a closed-form likelihood function, simulation-based estimation methods are required. This approach has been applied to panel datasets in several published research. Gonzalez et al. (2016) used panel random effect discrete choice to examine the change in travel time savings in Tenerife, Spain, before or after the deployment of streetcars. Jensen et al. (2013) study how individual preferences for adopting electric and conventional automobiles are influenced by experience. Using a panel rank-ordered mixed logit model, Srinivasan et al. (2006) were able to capture the relationship between better security levels and trip time increase in individuals' intercity mode choice. Reviewing the transportation research applications of panel random effect models is beyond the interest of this study.

On the other hand, we rather shift our focus to determine which features these applications share in adopting a random panel approach. Consequently, here we present a standard implemented structure for determining the utility function in random effect panel choice datasets. In equation (1), $I_{ni}$ is called the individual's $(n)$ intrinsic preference towards the alternative $(i)$. In random effect panel choice analysis, it is commonly believed that this term follows a normal or lognormal distribution. The term $\gamma_{ci,ni,t-1}$ in an individual's utility function carries over the individual's decision in the previous cycle to capture the choice persistence. However, this is rather a deterministic approach for capturing the persistence of choices, as it will be discussed at the end of this section, where RI-MNL is compared to these methods. The error term $\varepsilon_{ni,t}$ captures the unobserved components by the given systematic utility, and depending on its distribution assumption, the final model could take the form of a logit or probit family.

$$U_{ni,t} = I_{ni} + x_{ni,t} \cdot \beta + \gamma_{ci,ni,t-1} + \varepsilon_{ni,t}$$

The third strategy in analyzing panel datasets is the application of Dynamic Discrete Choice Models (DDCM) (Rust, 1987) in which decision-makers face repeated similar choice environments over time and optimize their intertemporal expected payoffs. In such models, time plays a prominent role, and the investigated research question is often the probability of choosing particular alternatives at a given time. Accordingly, implementations of this methodology frequently analyze panel datasets to explore the timing of an event's occurrence. For example, considerable research has used DDCM to study the number and timing of car ownership choices (Cirillo et al., 2016; Glerum et al., 2013; Liu and Cirillo, 2018). Another study utilized DDCM to estimate the evacuation behaviour of individuals over time during hurricanes (Rambha et al., 2021). Outside transportation research, DDCM is applied to various contexts, such as describing fertility behaviour, the quantity and time of having children (Wolpin, 1984), and examining the timing of construction choices in the housing market (Murphy, 2018). Reviewing the specifics of DDCM is beyond the scope of this study; interested readers are recommended to read Cirillo and Xu's (2011) review paper. All in all, compared to the two static approaches discussed earlier, DDCM approaches emphasize the temporal dimension of choice behaviour.

The RI-MNL panel model, which will be presented in the fourth section of this paper, falls between the second and third approaches and employs a Markovian process. Unlike random effect panel analysis, RI-MNL considers heterogeneity based on the variation of cost per unit of
information processing among individuals. In addition, because decision-makers are Bayesian agents, state dependence is an inherent characteristic of RI-MNL models. A benefit of RI-MNL panel analysis over random effect panel analysis is the consideration of preference persistence over choice persistence. Therefore, unlike the random effect model, which focuses on the persistence of selected alternatives, the prior information about the unselected alternatives is likewise carried forward to subsequent states. Last but not least, the RI-MNL panel model has a closed-form solution, making model estimation simpler than the random effect panel and dynamic discrete choice modelling.

3. Rational Inattention Multinomial Logit Model (RI-MNL)
The rational inattention theory assumes that decision-makers are rational. Still, since they do not have perfect information on the attributes of the alternatives in their choice set, they enter the decision-making process with two goals: maximizing their expected payoffs while minimizing the cost of information processing. The trade-off in this situation is that the more time agents spend processing information, the more confidence they gain in the expected payoffs, yet this confidence comes at the cost of processing information signals. This trade-off could be described as maximizing the following objective function:

$$\text{Max } [\text{expected value of payoffs} - \text{total information processing cost}]$$

(2)

To transform equation (2) into a mathematical problem, rational inattention theory assumes that decision makers are Bayesian agents who progressively update their knowledge of the payoffs of the alternatives in their choice set. Assume agent n has a prior belief $G(\vec{v})$ of the alternatives in their choice set for a single information acquisition process. Then, if this agent receives a new signal of information $\vec{s}$, they can form their posterior belief of the payoff function conditional on the acquired signal of information (equation 3) (The index $n$ is dropped from the equations for simplicity).

$$f(\vec{v}|\vec{s}) \propto G(\vec{v}) \cdot f(\vec{s}|\vec{v})$$

(3)

The posterior belief conditional on a single information signal $\vec{s}$ is demonstrated by Equation (3). Throughout the process of acquiring information, decision-makers get several signals of information on each alternative $i$ to update their knowledge of posterior beliefs. The posterior belief can be expressed as the Lebesgue integration shown below. The posterior belief can be expressed as the Lebesgue integration in equation (4).

$$P_i(\vec{v}) = \int f(d\vec{s}|\vec{v})$$

(4)

After each cycle of information processing, decision-makers form the unconditional probability of alternative $i$ and revise their previous beliefs:

$$P^0_i = \int P_i(\vec{v}) \cdot G(d\vec{v})$$

(5)
The expected payoff for each alternative $i$ is the payoff for that alternative weighted by the unconditional probability of selecting that alternative. As the choice has not yet been made at the stage of the information acquisition process, the expected value of payoffs is the summation of each alternative's expected value $i$:

$$

\text{Expected payoffs} = \sum_i \int \bar{v} \cdot P_i(\bar{v}) \cdot G(d\bar{v})
$$

(6)

Using equation (6) and the objective function in equation (2), the mathematical model in equation (7) could be derived:

$$

\text{Max } \sum_i \int \bar{v} \cdot P_i(\bar{v}) \cdot G(d\bar{v}) - \hat{C}(f)
$$

subject to:

$$

P_i(\bar{v}) \geq 0 ; \forall \bar{v} \in \mathbb{R}^N \quad (7-1)
$$

$$

\sum_{i=1}^N P_i(\bar{v}) = 1 ; \forall \bar{v} \in \mathbb{R}^N \quad (7-2)
$$

The term $\hat{C}(f)$ in optimization problem (7) is a function of information processing cost that is dependent on the function of information signals $f(\bar{s}|\bar{v})$. The objective function in (7) must be solved under two conditions to ensure that probabilities are positive (7-1) and their sum across each signal state is equal to one (7-2).

The intuition behind the optimization problem (7) is straightforward. In RI theory, decision-makers also gain "utility" from the alternatives they do not choose, as long as they acquire information about the payoffs of those alternatives (One can think of this as assurance for knowing that alternatives that are not selected would not have provided more payoff to the user). However, the cost of processing information $\hat{C}(f)$ imposes a cognitive barrier on their ability to collect perfect information on all alternatives in their choice set.

In the RI-MNL model, the information processing cost $\hat{C}(f)$ is identified as Shannon's entropy (Shannon, 1948). Shannon's entropy states that less uncertainty always leads to a decrease in entropy $H$. Therefore, the state of entropy of the decision maker before and after receiving information signals is different, and the acquired information may be measured using their difference in equation (8):

$$

I(\bar{v}, \bar{s}) = H[f(\bar{v})] - H[f(\bar{v}|\bar{s})]
$$

(8)

Shannon argues that since rare incidents release more information, the entropy could be stated as an inverse logarithm of probabilities. Consequently, equation (8) could be expanded to equation (9), where the difference between the expected released information for the two states is calculated.
\[ I(\vec{v}, \vec{s}) = \int \left( \sum_{i=1}^{N} P_i(\vec{v}) \cdot \log P_i(\vec{v}) \right) \cdot G(\,d\vec{v}) - \sum_{i=1}^{N} P_i^0 \cdot \log P_i^0 \]  
\hspace{2cm} (9)

The information obtained through information signals is measured in "bits" by equation (9). Because \( \hat{C}(f) \) in the mathematical model, (7) must be in "utils," equation (9) must be multiplied by a term called the unit of information processing cost, which is measured in "utils/bits." As a result, the final form of \( \hat{C}(f) \) is:

\[ \hat{C}(f) = \lambda \cdot \left( \int \left( \sum_{i=1}^{N} P_i(\vec{v}) \cdot \log P_i(\vec{v}) \right) \cdot G(\,d\vec{v}) - \sum_{i=1}^{N} P_i^0 \cdot \log P_i^0 \right) \]  
\hspace{2cm} (10)

The mathematical model (7) can be transformed to the following Lagrangian function with all of its constraints:

\[ \mathcal{L}(P) = \sum_{i} \int \vec{v}_i \cdot P_i(\vec{v}) \cdot G(\,d\vec{v}) - \lambda \left[ \int \left( \sum_{i=1}^{N} P_i(\vec{v}) \cdot \log P_i(\vec{v}) \right) \cdot G(\,d\vec{v}) - \sum_{i=1}^{N} P_i^0 \cdot \log P_i^0 \right] + \int \gamma(\vec{v}) \cdot P_i(\vec{v}) \cdot G(\,d\vec{v}) - \int \mu(\vec{v}) \cdot \left( \sum_{i=1}^{N} P_i(\vec{v}) - 1 \right) \cdot G(\,d\vec{v}) \]  
\hspace{2cm} (11)

Where \( \gamma(\vec{v}) \) and \( \mu(\vec{v}) \) are the Lagrangian multipliers on conditions (7-1) and (7-2). The first derivative of the \( \mathcal{L}(P) \) with respect to \( P_i(\vec{v}) \) is then equal to:

\[ \mathcal{L}'(P) = \vec{v}_i + \lambda \cdot (\log P_i^0 + 1 - \log P_i(\vec{v}) - 1) + \gamma(\vec{v}) - \mu(\vec{v}) \]  
\hspace{2cm} (12)

Matějka and McKay (2015), in deriving the RI-MNL model, prove that in optimizing equation (12), condition (7-1) is a non-binding condition. Therefore, optimizing \( \mathcal{L}'(P) = 0 \) results in:

\[ P_i(\vec{v}) = P_i^0 \cdot e^{(\vec{v}_i - \mu(\vec{v}))/\lambda} \]  
\hspace{2cm} (13)

Then plugging (13) into (7-2) will result in:

\[ \sum_{i=1}^{N} P_i^0 \cdot e^{\vec{v}_i/\lambda} = e^{\mu(\vec{v})/\lambda} \]  
\hspace{2cm} (14)
Finally, the probability of the selection of each alternative is derived by replacing the $e^{\mu(\bar{v})/\lambda}$ term in equation (13) in equation (14):

$$P_t(\bar{v}) = \frac{P_i^0 e^{\nu_i/\lambda}}{\sum_{j=1}^J P_j^0 e^{\nu_j/\lambda}}$$

(15)

Because of its resemblance to the classical multinomial logit model, equation (15) is referred to as the RI-MNL model. However, there are two key distinctions between the RI-MNL model and the standard MNL model:

1- The unconditional probability $P_i^0$ in the RI-MNL model, which comes from equation (5), represents the decision maker's prior view and must be identified separately. The RI-MNL model could be either closed or non-closed, depending on the $P_i^0$ specification.

2- Individuals' information processing costs $\lambda_n$ must be identified separately in the RI-MNL model. The identifications need to consider that $\lambda_n$ is the source of taste variation in choice behaviour. In the standard MNL model, however, $\lambda$ is the scale parameter of the Extreme Value Type I distribution and is not identified separately.

To summarize, both the prior unconditional probability $P_i^0$ and the information processing cost unit $\lambda_n$ must be identified to establish an empirical specification for the RI-MNL model. The following section of this study proposes a set of assumptions that aid the identification of $P_i^0$ and $\lambda_n$ terms for in the RI-MNL model in panel choice datasets.

**4. RI-MNL Estimation for Panel Dataset**

Any empirical RI-MNL model specification must identify the unconditional probability $P_i^0$ and the unit of information processing cost $\lambda$. This section presents the identification of unconditional probability $P_i^0$ using a Markovian process and a linear-in-parameter information processing cost function for $\lambda$ identification. This section will begin by explaining the Markovian process, followed by a discussion of the assumptions underlying the linear function measurement of unit information processing cost.

Regarding the identification of unconditional probabilities, we assume that each cycle of the panel dataset wherein decision behaviour is recorded corresponds to a stage in the Bayesian process through which individuals update their knowledge of the payoffs of alternatives in their choice set. For example, if a panel dataset is collected in four cycles, the information available for Bayesian agents updating their posterior knowledge is limited to four stages of data collection. This assumption may not be an exact reflection of what happens in reality. During the period covered by the panel dataset, individuals may move through several states of updating their knowledge of the payoff value of alternatives. However, "processing acquired information signals" is more of an abstract concept, making it nearly impossible to attempt to measure this process directly from the collected datasets. As a result, the RI-MNL model specification's challenge is that individuals go through an information acquisition and processing phase that is not observable by the researcher. Theoretically, this implies that equations (4) and (5) in section three are immeasurable and unconditional probabilities could not be identified directly from these measurements.
The previously stated assumptions aid in overcoming this problem and identifying unconditional probability indirectly from the results (i.e., choice behaviour) of the information acquisition and processing phase. In other words, one can effectively measure the changes in unconditional probability if they observe serial changes in choice behaviour across the many cycles of a panel dataset. Consider a panel dataset with \( k \) cycles. The unconditional probability \( p_{i,tk}^{0} \) represents the unconditional probability for alternative \( i \) at the time step \( t_k \). To indirectly measure unconditional probabilities, we assume that the unconditional probability at the time step \( t_k \) is measurable via the unconditional probability at the time step \( t_{k-1} \) in a Markovian process using a RI-MNL model (see figure 1).

\[
p_{i,tk}^{0} = \frac{p_{i,t_{k-1}}^{0}e^{v_{i,tk}/\lambda}}{\sum_{j=1}^{J} p_{i,t_{k-1}}^{0}e^{v_{j,tk}/\lambda}} \tag{16}
\]

As shown in figure 1, the unconditional probability of each state is calculated using the RI-MNL model, with the unconditional probability of the preceding state as an input. Therefore, the probabilities of each alternative's selection \( P_i(\vec{v}) \) might be determined using simply the unconditional probability of the final state (equation 17).

\[
P_i(\vec{v}) = \frac{p_{i,t_K}^{0}e^{v_{i,t_K}/\lambda}}{\sum_{j=1}^{J} p_{j,t_K}^{0}e^{v_{j,t_K}/\lambda}} \tag{17}
\]

In other words, the proposed Markovian process reduces the identification of the unconditional probability \( p_i^{0} \) to the modellers' initial value assumption for \( p_{i,0}^{0} \). Using the data acquired during the many cycles of the dataset, the value of \( p_{i,tk}^{0} \) is then updated in each state to create a more accurate estimate of the individual's \( p_i^{0} \).

Another point worth mentioning in relation to figure 1 is the time-dependent definition of the unit of information cost in the proposed Markovian process. Depending on the context, defining an evolving function for the information processing cost unit could be critical. For example, in
the context of long-term decisions, it is important to consider a number of factors whose occurrence will alter an individual's attentiveness (e.g., a change in marital status for the residential mobility decision), whereas, in the context of short-term decisions, an individual's attentiveness may decrease over time as a result of the habit of making repeated decisions (e.g. Changes in the likelihood of trying a new daily trip route for individuals who repeatedly choose the same route out of habit).

The proposed Markovian process addresses the identification of unconditional probability in panel choice datasets. For the RI-MNL model to be fully specified, the unit of information processing cost must also be determined. As previously noted, information processing is an abstract concept, making it difficult to identify a strategy for absolute measurement of information processing cost unit \( \lambda \). In theory, the unit of information processing cost should represent how attentive individuals are in decision behaviour, and it should also be useful in discovering sources of variation in choice behaviour. Similar to the method described to determine unconditional probability, this study presents an indirect method for measuring the unit of information processing cost \( \lambda \) by tracing the heterogeneity in choice behaviour.

According to rational inattention theory, the unit of information processing cost \( \lambda \) could take any positive value. Small values of \( \lambda \) indicate that a person has a low information processing cost and, accordingly, receives fewer "utils" when collecting "bits" of information signals. Therefore, such individuals demand more information signals, which increases the likelihood that their posterior beliefs would differ from their prior beliefs. In this study, individuals exhibiting this type of choice behaviour are considered "rational perfectionist consumers." On the opposite end of the spectrum, individuals with large values of \( \lambda \) who acquire a substantial amount of "utils" by receiving a single "bit" of information. Such individuals find information processing tasks overwhelming and become insensitive to incoming signals. In this study, these individuals are referred to as "rational inattentive consumers," and they often display inertia towards their priors in their decision-making behaviour.

With the discussions above, the \( \lambda \) identification challenge is condensed to determining the perfectionist and inattentive behaviour in the dataset. If one considers the two extremes of the attentiveness continuum, one will find fully inattentive consumers on one end and fully perfectionist consumers on the other. A fully inattentive consumer is a person who is nonresponsive to new information signals, and it is highly improbable that they will make a decision that is contrary to their previous choices. On the other side, fully perfectionist consumers are always looking for a better option. Bringing this argument to an aggregate perspective, it is reasonable to assume that the diversity in choice behaviour is correlated with attentiveness, where perfectionists have larger \( \lambda \) values, and lower \( \lambda \) values correspond to inattentive choice behaviour. Our strategy for identifying \( \lambda \) is founded on this reasoning. Thus, the unit of information processing cost \( \lambda \) in this empirical study is described as a function of parameters to represent heterogeneity in decision-making. The selection of these parameters should be based on the factors that experts believe could be the source of behavioural heterogeneity.

\[ y_n = y_n \cdot z_n \quad (18) \]
Equation (18) is a linear model of the vector of selected factors $\mathbf{z}_n$ and $\mathbf{y}_n$ as their corresponding parameters. The $y_n$ estimator is subject to a constraint, as it is not defined for negative values. To ensure that the cost per unit of data processing $\lambda_n$ remains positive, we propose the following transformation (figure 2):

$$\lambda_n = 1 + \tanh(y_n)$$ (19)

The transformation formulation in equation (19) assures that the cost of processing information is always positive. This transformation is chosen because it possesses two key properties: 1) This transformation has a central symmetry with regard to (0,1), which is set to be the reference value for the unit of information processing cost ($\lambda_0 = 1$); 2) This transformation makes numerical maximum likelihood estimation more stable by preventing extreme values of $\lambda_n$ resulting from data outliers. In equations (18) and (19), it is crucial to note that $\lambda_n$ represents the unit of information processing cost for individual $n$. In part three, the index $n$ was omitted to simplify the derivation of the RI-MNL model. Finally, with the identifications stated in this section for the panel datasets, the RI-MNL model's structure is as follows:

$$P_{in}(\vec{v}) = \frac{P_{in}^{0,tK-1} e^{v_{in,tK}/(1+\tanh(y_nz_n))}}{\sum_{j=1}^{J} P_{jn}^{0,tK-1} e^{v_{jn,tK}/(1+\tanh(y_nz_n))}}$$ (20)

The closed form panel RI-MNL model can be estimated using maximum likelihood estimation (equation (20)). Since the suggested RI-MNL model employs a Markovian process, it will include the simultaneous estimation of at least two RI-MNL models. In the case of a two-cycle panel dataset, for example, the likelihood function of the panel RI-MNL model has the following form:

$$\mathcal{L}(\beta \mid x_{t1}, x_{t2}) = \left\{ \prod_{i} \mathcal{L}_1(\beta_1 \mid x_{t1}) \right\} \times \left\{ \prod_{i} \mathcal{L}_2(\beta_2 \mid x_{t2}) \right\}$$ (21)
The proposed model specification is applied to a panel residential relocation stated preference dataset in the next section to illustrate the panel RI-MNL application. The Aptech GAUSS MAXLIK MT 2.0 application package was used to estimate the empirical model in this study.

5. Empirical Analysis

For the empirical investigation of the panel RI-MNL model, this study uses a panel stated preference dataset containing residential location preferences. The purpose of the implemented survey was to examine the impact of the COVID-19 pandemic on Greater Toronto Area residential relocation behaviour. The collection of data for this dataset occurred in two cycles in July 2020 and 2021. In the choice experiments conducted in both cycles, respondents were informed of the dwelling characteristics and location of four new options, as well as the pandemic status and telecommuting conditions, and then asked whether they would relocate to one of the four given dwellings. In parallel with the report on the best-fitting model, the attributes and levels of choice experiments will be discussed in further detail.

For empirical analysis, 158 respondents from both dataset cycles were separated. Each participant has completed nine multiple-choice tasks. In each task, respondents were given the number of days they could telecommute and asked to relocate their residence under three different pandemic situations. The model structure proposed in section four of this paper is applied to this dataset, and multiple model structures were examined. The model with the highest rho-squared value and the most intuitive outcomes is chosen as the best. For the purpose of evaluating the effectiveness of the proposed panel RI-MNL model, we estimated a standard MNL model and multiplied likelihoods across the same choices made by each individual to account for panel effects in the dataset. Our findings show that the panel RI-MNL model has a higher rho-squared value (0.2286) than the panel MNL model (0.1874).

This section's estimation results are separated into two parts. The parameters for the payoff function are presented in Table 1. The findings of the information processing cost function are then provided in Table 2.

| Table 1- Panel RI-MNL Model Results - Payoff Estimations |
|----------------------------------------------------------|
| **Panel cycle** | **Sub-model** | **Final Model** |
|                | **Cycle I (2020)** | **Cycle II (2021)** |
| **Region variables:** | **Estimations** | **Estimations** |
| Ajax           | 1.216 | 6.716 | <0.001 | 1.591 | 4.191 | <0.001 |
| Brampton       | 0.575 | 2.999 | 0.003 | 0.472 | 1.586 | 0.113 |
| Downtown       | 0.018 | 0.102 | 0.919 | 2.368 | 7.925 | <0.001 |
| East End Toronto | 1.221 | 8.539 | <0.001 | 1.125 | 4.201 | <0.001 |
| Etobicoke      | 1.686 | 13.068 | <0.001 | 0.523 | 2.071 | 0.038 |
| Halton Hills (reference dummy cycle II) | --- | --- | --- | 0 | --- | --- |
| Markham        | 0.752 | 5.022 | <0.001 | 1.688 | 6.493 | <0.001 |
| Milton (reference dummy both cycles) | 0 | --- | --- | 0 | --- | --- |
| Neighbourhood | Coefficient | t-value | P-value |
|---------------|-------------|---------|---------|
| Detached Detached variable | 0.374 | 9.163 | <0.001 |
| Semi-detached | 0.051 | 1.041 | 0.298 |
| Condo (reference dummy) | 0 | --- | --- |
| Townhouse | -0.096 | -2.07 | 0.040 |
| Price change | -0.064 | -3.1 | 0.002 |
| Area change | 0.04 | 1.835 | 0.067 |
| Neighbourhood Quality (shared) reference dummy: busy and noisy | 0 | --- | --- |
| no relocation: moderate noise | 0.509 | 5.778 | <0.001 |
| no relocation: green and quiet | 1.124 | 8.576 | <0.001 |
| relocation: moderate noise | 0.253 | 1.913 | 0.058 |
| relocation: green and quiet | 0.895 | 7.318 | <0.001 |
| Transit Accessibility reference dummy: Not easily accessible | 0 | --- | --- |
| no relocation: moderate access | 0.533 | 4.434 | <0.001 |
| no relocation: quick access | 0.551 | 3.548 | <0.001 |
| relocation: moderate access | 0.066 | 1.038 | 0.301 |
| relocation: quick access | 0.088 | 1.152 | 0.251 |
| Highway Accessibility reference dummy: Not easily accessible | 0 | --- | --- |
| no relocation: moderate access | 0.217 | 2.231 | 0.027 |
| no relocation: immediate access | 0.524 | 5.154 | <0.001 |
| relocation: moderate access | -0.046 | -0.628 | 0.531 |
| relocation: immediate access | 0.351 | 3.182 | <0.001 |
As stated in the previous section, the final specification of the panel RI-MNL model requires the researcher to assume the initial states preceding the first cycle in the panel data to be 0,0. Since no comparable choice dataset was gathered before the COVID-19 pandemic, we assumed uniform distribution for 0,0 to begin this empirical research. This is akin to assuming no prior information regarding the choice behaviour (equiprobable choices).

Before the interpretation of estimated payoff variables (table 1), two key points must be made. First, table 1 results are structured to list the included SP attributes alongside the estimated RI-MNL model coefficients. Second, the parameters in the payoff function for the no relocation decision are separated from those in the payoff functions for the four offered relocation decisions. We tested several modelling structures and determined that this division is important to provide a better model fit. The existence of individuals' inertia to remain in their current residence is a rational explanation for the necessity of this separation. In general, our findings indicate that individuals derive greater benefits from their current residence if it shares a characteristic with one of the provided dwellings.
The parameters for the dwelling types are fixed between the sub-model and the final model. This decision was made to stabilize model estimation and guarantee that the parameters of both models remain within a close range. The estimation results demonstrate a general trend toward detached dwelling types. Moreover, model estimates indicate that townhouses are the least popular dwelling type in the region.

The model includes the region attribute to capture respondents' intrinsic preference for a particular region. The reference dummies are purposefully set to regions that make other coefficients positive to facilitate comparisons. Since the Halton Hills region was only added in the second cycle of this study, Milton was grouped with Halton Hills as the reference dummy variable for the second cycle. We validated the assumption of combining (Milton and Halton Hills) by applying a separate model for cycle two and determining that the difference between Halton Hills and Milton region parameters was not statistically significant. The most intriguing discovery regarding region parameters is the significant shift in attitude towards the downtown Toronto region between the two cycles. This finding demonstrates that the COVID-19 vaccination and province reopening phases successfully increased demand for the downtown Toronto region after the lockdown phase in cycle 1.

The payoff functions incorporated price and area attributes relative to the value and area of the current residence of an individual. For both variables, the sign and magnitude of coefficients are reasonable, and there are no discernible differences between the final and sub-model values.

The attributes that fortified the separation of coefficients for no relocation and relocation payoff functions were neighbourhood quality, transit and highway accessibility, and neighbourhood biking paths. The results for these attributes indicate two major relocation preference trends. First, households gain a greater payoff for comparable attribute levels when they choose not to relocate as opposed to the relocation decision. This observation indicates that households prefer to keep what they have rather than risk losing it by attempting to obtain something better. The second significant trend relates to the relocation payoffs, wherein households only target the highest level of each attribute when they make the relocation decision. In other words, households are unwilling to relocate unless they relocate to a residence that provides the highest level of the desired attribute.

The findings regarding telecommuting and the pandemic conditions of COVID-19 are thought-provoking. These two characteristics were only included in no-relocation payoff functions, and their positive values indicate a reduced tendency for relocation and vice versa. Based on coefficient comparisons between the first and second cycles, it appears that during the second cycle, individuals became insensitive to the pandemic status of COVID-19. Our sub-model (first cycle) findings indicate that the COVID-19 pandemic halts relocation decisions, and individuals avoid making decisions due to the pandemic's uncertainty. However, this conclusion is reverted for the second cycle, and the decision regarding residential relocation appears to be independent of the pandemic status.

The number of days available for telecommuting is the final payoff attribute to be discussed. The findings show that people consider relocating if they are offered more telecommuting days. Comparing the two cycles reveals that the value gained per offered day of telecommuting has
doubled between 2020 and 2021. This finding may be indicative of an increasing attitude towards telecommuting among Greater Toronto Area employees.

Table 2- Panel RI-MNL Model Results - Information Processing Cost Function Estimations

| Variable                              | Estimate | t-Value | P-Value |
|---------------------------------------|----------|---------|---------|
| Gender: male                          | 0.177    | 2.464   | 0.015   |
| Household type: couple no children    | -0.162   | 3.419   | <0.001  |
| Household type: single individuals    | -0.419   | 3.808   | <0.001  |
| Log (years at the current residence)  | 0.359    | 2.313   | 0.022   |
| Log (age)                             | 0.333    | 3.014   | 0.003   |

- Household type entered the model as dummy variables, and the dummy reference is "Family with children."

The estimation results for the information processing cost function are presented in Table 2. These findings are useful for associating choice behaviours with the attentiveness of individuals. Based on the results of this empirical model, the least attentive individuals for residential relocation decisions are elderly males who reside in a family with children and who have lived in their current residence for a long time. Young, single females who have only lived in their current residence for a few years are the most flexible regarding residential relocation. In addition to the variables shown in table 2, marital status, employment status, and education level were tested in the information cost processing function and found to be insignificant or highly correlated with other explanatory variables.

An intuitive finding of the information processing cost function presented in table 2 is that these variables could alter individuals along the attentiveness spectrum throughout the panel cycles. For instance, those who continue to reside in their current residence will pay less attention to future relocation decisions. Despite this, the identified information processing cost function cannot be considered ideal because it lacks data on major life events. It is widely argued that the decision to relocate is tied to major life events such as a job change, childbirth, or a family member's death. However, the mentioned information was not present in the collected panel dataset for inclusion in the information processing cost function.

6. Conclusion for Future Research Direction

The rational inattention theory examines the behaviour of Bayesian agents in an environment rich in information that consumes their attention. There has been a recent shift in interest in applying this theory to the context of discrete choice models; however, these advances have been primarily theoretical. This study contributes by proposing an empirical specification for panel RI-MNL that employs a Markovian process to identify prior unconditional probabilities and a linear function to measure the information processing cost unit.

We identify three principal advantages of applying rational inattention theory to panel choice datasets:

- The assumption that decision makers are Bayesian agents inherently generates a framework that considers state dependence and inertia towards previous decisions.
The variation in choice behaviour could be linked with selected variables, which is more intuitive than the standard practice of capturing taste heterogeneity through random component models.

Because of the closed form formulation, estimating the RI-MNL model on panel data requires less computational effort compared to random effect and dynamic discrete choice models, making it more attractive to implement. However, this implementation is hindered by the lack of pre-programmed application packages and software for RI-MNL model estimation, requiring researchers to code the model estimation themselves.

To investigate the effects of the COVID-19 pandemic on residential relocation preferences, the proposed panel RI-MNL model was applied to a panel dataset collected in two cycles. The results indicate that as the pandemic progresses, respondents become less sensitive to its status. On the other hand, the overall attitude towards telecommuting has improved, which may indicate a growing social acceptance of telecommuting, as respondents are beginning to consider this factor when making long-term decisions, such as relocating their residence.

The empirical panel RI-MNL model presented in this study has some limitations. This study associates attentiveness with a limited set of variables that presumably cannot fully capture the attentiveness of respondents. Major life events frequently influence residential relocation decisions. However, these potential occurrences are not included in the model due to a lack of data. A further limitation of this study is the assumption of equiprobable market shares for the prior beliefs in the first cycle, as there was no comparable study conducted in the region that could be used as a reference for determining the initial values for the unconditional probabilities of the sub-model.
REFERENCES

Ambos TC, Cesinger B and Eggers F (2020) How does de‐globalization affect location decisions? A study of managerial perceptions of risk and return. Global Strategy Journal 10(1). Wiley-Blackwell Publishing Ltd: 210–236. DOI: 10.1002/gsj.1335.

Cirillo C and Xu R (2011) Dynamic discrete choice models for transportation. Transport Reviews 31(4): 473–494. DOI: 10.1080/01441647.2010.533393.

Cirillo C, Xu R and Bastin F (2016) A dynamic formulation for car ownership modeling. Transportation Science 50(1): 322–335. DOI: 10.1287/trsc.2015.0597.

Fosgerau M, Melo E, de Palma A, et al. (2020) Discrete Choice and Rational Inattention: A General Equivalence Result. International Economic Review 61(4). John Wiley & Sons, Ltd: 1569–1589. DOI: 10.1111/IERE.12469.

Glerum A, Frejinger E, Karlström A, et al. (2013) A dynamic discrete-continuous choice model of car ownership, usage and fuel type. In: Second Symposium of the European Association for Research in Transportation, 2013. Available at: https://infoscience.epfl.ch/record/195888.

González RM, Marrero ÁS and Marrero GA (2016) How the values of travel time change when a panel data around a new tram implementation is used. European Journal of Transport and Infrastructure Research 16(4). Editorial Board EJTIR: 554–572. DOI: 10.18757/EJTIR.2016.16.4.3158.

Hanly M and Dargay JM (2000) Car Ownership in Great Britain: Panel Data Analysis: Transportation Research Record (1718): 83–89. DOI: 10.3141/1718-11.

Jensen AF, Cherchi E and Mabit SL (2013) On the stability of preferences and attitudes before and after experiencing an electric vehicle. Transportation Research Part D: Transport and Environment 25. Pergamon: 24–32. DOI: 10.1016/J.TRD.2013.07.006.

Jensen AF, Cherchi E and de Dios Ortúzar J (2014) A long panel survey to elicit variation in preferences and attitudes in the choice of electric vehicles. Transportation 41(5). Springer New York LLC: 973–993. DOI: 10.1007/S11116-014-9517-6.

Liu Y and Cirillo C (2018) A generalized dynamic discrete choice model for green vehicle adoption. Transportation Research Part A: Policy and Practice 114. Elsevier Ltd: 288–302. DOI: 10.1016/j.tra.2018.01.034.

Matějka F and McKay A (2015) Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model. American Economic Review 105(1). American Economic Association: 272–98. DOI: 10.1257/AER.20130047.

McFadden D (1973) Conditional Logit Analysis of Qualitative Choice Behavior. Institute of Urban and Regional Development, University of California.

McFadden D (1978) Modeling the choice of residential location. Transportation Research Record (673).

McFadden D and Train K (2000) Mixed MNL models for discrete response. Journal of applied Econometrics 15(5). Wiley Online Library: 447–470.

Meeran S, Jahanbin S, Goodwin P, et al. (2017) When do changes in consumer preferences make forecasts from choice-based conjoint models unreliable? European Journal of Operational Research 258(2). North-Holland: 512–524. DOI: 10.1016/J.EJOR.2016.08.047.
Murphy A (2018) A dynamic model of housing supply. American Economic Journal: Economic Policy 10(4). American Economic Association: 243–267. DOI: 10.1257/POL.20150297.

Rambha T, Nozick LK and Davidson R (2021) Modeling hurricane evacuation behavior using a dynamic discrete choice framework. Transportation Research Part B: Methodological 150. Elsevier Ltd: 75–100. DOI: 10.1016/j.trb.2021.06.003.

Rust J (1987) Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. Econometrica: Journal of the Econometric: 999–1033. Available at: https://www.jstor.org/stable/1911259.

Shannon CE (1948) A mathematical theory of communication. The Bell system technical journal 27(3). Nokia Bell Labs: 379–423.

Simon HA (1984) Models of bounded rationality, volume 1: economic analysis and public policy. MIT Press Books 1. The MIT Press.

Sims CA (2003) Implications of rational inattention. Journal of Monetary Economics 50(3). North-Holland: 665–690. DOI: 10.1016/S0304-3932(03)00029-1.

Srinivasan S, Bhat CR and Holguin-Veras J (2006) Empirical Analysis of the Impact of Security Perception on Intercity Mode Choice: A Panel Rank-Ordered Mixed Logit Model. Transportation Research Board 1942(1): 9–15. DOI: 10.1177/0361198106194200102.

Wen CH and Koppelman FS (2001) The generalized nested logit model. Transportation Research Part B: Methodological 35(7): 627–641. DOI: 10.1016/S0191-2615(00)00045-X.

Wolpin KI (1984) An Estimable Dynamic Stochastic Model of Fertility and Child Mortality. Journal of Political Economy 92(5). University of Chicago Press: 852–874. DOI: 10.1086/261262.