# Dynamic Choice Models

Michel Bierlaire * Emma Frejinger † Tim Hillel *

March 5, 2021

Report TRANSP-OR 210305
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
transp-or.epfl.ch

*École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory, Switzerland, {michel.bierlaire,tim.hillel}@epfl.ch

†Université de Montréal, Department of Computer Science and Operations Research, Canada, emma.frejinger@umontreal.ca

# 1 Introduction

Real-world choice situations are often *dynamic* - choices made in the present are dependent on choices made in the past and, in turn, will also affect future choices. For the sake of illustration, consider the simple example of a student who is given a weekly budget to purchase her lunch in the school canteen. The choice of which option to take for lunch on any one day is dependent on the student's remaining budget, which itself is dependent on the purchases she has made up until that point. Furthermore, the student's perceptions of each available option may also be dependent on her previous choices. For example, the student may learn over time which types of food in the canteen tend to be higher quality. Finally, the student might include future planning in her decision process. For example, she may choose a lower cost option one day to ensure she has enough budget for her favourite (and more-expensive) option which tends to be offered on a later day in the week.

Dynamic choice models are a family of models that describe sequential choices (such as those described in the above example) by attempting to capture changes in the decision process over time. Estimating these models therefore requires *panel* data, which provides details of multiple sequential choices of individuals over time. There are many possible mechanisms for dynamic behaviour, each of which may be included (or not) in different modelling scenarios. The situation described above presents three; (i) *changes in external factors over time*, (ii) *habitual behaviour and learning*, and (iii) *forward-looking planning*.

There are a plethora of other situations where individuals are faced with sequential choices over extended periods of time. Examples include: car ownership decisions (there is an extensive literature on this topic, see, Cirillo et al., 2015, for a survey); retirement planning (Rust & Phelan, 1997); and career decisions (Keane & Wolpin, 1997). Furthermore, certain choice situations that take place over relatively short periods of time can be naturally formalized as sequential decision-making problems. Route choice is such an example (Zimmermann & Frejinger, 2020) and we use it for the sake of illustration in the following.

Consider an individual choosing a path in a network composed of a set of nodes and a set of arcs. Here we consider the case when the network represents a road network where each node corresponds to an intersection and each arc to a road segment (Fosgerau et al., 2013). We note that the network could also be an abstract representation of many different choices. Examples include daily transportation mode and activity choices (Västberg et al., 2019) and location choice (Danalet et al., 2016). An individual's choice of path between a given origin-destination pair can be decomposed into a sequence of arc choices, where, starting at the origin and at each intersection, the individual chooses the next road segment. While making the choice of road segment, the individual is forward-looking as she seeks to reach the destination. If she is perfectly forward-looking, then a sequential choice model can be equivalent to a non-sequential (path-based) one (Fosgerau et al., 2013). However, the sequential model presents a number of advantages over path-based approaches, in particular from a computational point of view.

In this chapter, we present a generalised formulation of the dynamic choice problem, and demonstrate how it encapsulates the three aforementioned mechanisms. This problem formulation is then used to derive a general parametric dynamic choice model which can be estimated from data. Finally, we show how different assumptions on our generalised parametric model can be used to derive various examples of dynamic choice models from the literature. This approach allows us to unify the diverse existing dynamic choice modelling approaches in the literature under a unified framework and illustrate the differences between models through their implied assumptions.

We use the following notation throughout the chapter. An individual $n$ makes choices within a set $\mathcal{C}$ of $J$ alternatives over a time horizon. The latter is discretised in several time intervals indexed by $t = 0, \dots, T$, not necessarily of equal length. The assumption is that all the variables involved in the process are constant within each time interval, but may vary from one interval to the next. The number of time intervals $(T + 1)$ is supposed to be finite. For the sake of notational simplicity and without loss of generality, the set $\mathcal{C}$ is assumed to be constant over $n$ and $t$, and contains every possible alternative that can be chosen by all individuals across all time intervals. The notation in this chapter obeys the following convention: (i) lower case letters refer to deterministic variables; (ii) upper case letters refer to random variables; and (iii) Greek letters are used to refer to model parameters and error terms.

The rest of this chapter is laid out as follows. In the following section, we outline the dynamic discrete choice problem from the point of view of the decision maker and introduce the utility maximization problem they solve at each time $t$. Section 3 then presents the same problem from the point of view of the analyst and introduces the dynamic programming formulation of the optimization problem faced by the decision maker. In Section 4, we specify a general parametric model and summarise how it can be estimated from historic data. Section 5 then introduces two different approaches to account for habitual behaviour and learning, based on the Markov assumption. Once the generalised parametric model has been established, Section 6 demonstrates how different types of dynamic choice models in the literature can be derived through applying specific assumptions on the parameters of the generalised model. Finally, Section 7 summarises the chapter and presents avenues for future research.

## 2 The point of view of the decision maker

At time interval $t$, the individual $n$ chooses a single alternative $i_{nt}$ in the choice set $\mathcal{C}$. The availability of each alternative for each individual at each time interval is characterized by binary variables $a_{int}$, with value 1 if alternative $i$ is available for individual $n$ at time $t$, and 0 otherwise[1]. The choice made by the individual is based on the knowledge acquired from the past as well as the anticipation of the impact of the choice on future outcomes. The decision variables are defined as

$$y_{int} = \begin{cases} 1 & \text{if } i = i_{nt}, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

We denote by $y_{nt} = y_{1:J,nt}$ the decision vector for individual $n$ at time $t$ and by $y_n = y_{n,0:T}$ the *trajectory*, that is, the sequence of decisions made by individual $n$ over time. It is useful to consider the set of the $J$ feasible decisions that any individual can take at any

[1]The inclusion of availability indicators $a$ allows for a constant choice set C (as specified in Section 1) without any loss of generality.

point in time. This is denoted by

$$\mathcal{Y} = \{\delta_i, i \in \mathcal{C}\}, \tag{2}$$

where $\delta_i$ is a vector of length $J$, such that all entries are zero, except entry $i$ that is 1. It characterizes the choice of alternative $i$. We can also consider the set of feasible trajectories, i.e., the set of feasible decisions that an individual can take during the whole horizon. This is obtained by considering the Cartesian product of $\mathcal{Y}$ over time:

$$\mathcal{T} = \bigtimes_{s=0}^{T} \mathcal{Y}. \tag{3}$$

The cardinality of $\mathcal{T}$ is $(T+1)^J$, which is usually too large to allow for an explicit enumeration of the set.[2] We also denote by

$$\mathcal{T}_t = \bigtimes_{s=t}^{T} \mathcal{Y} \tag{4}$$

the set of trajectories starting at time $t$.

The data that is available to the decision maker at time interval $t$ to perform her choice is represented by the vector $\tilde{x}_{nt}$. This vector contains all information the decision maker uses to make her decision, including attributes of each alternative in the choice set (possibly including their historical values), as well as her previous choices and outcomes (i.e. utility functions). This allows for habits and learning to be captured, as discussed in Section 5. The vector also includes context variables, that may vary over time (e.g., weather), and the availability indicators $a_{int}$. Note that the vector $\tilde{x}_{nt}$ may contain both discrete and continuous variables. However, in order to simplify the formulations below, we systematically use integrals and density functions, as if all explanatory variables were continuous.

The individual may also anticipate the impact of her decision on the future values of the explanatory variables. As such, if $t$ is the current time interval, then for a future interval $s > t$, $\tilde{X}_{ns}(t)$ is a vector of random variables. which represents the individual's anticipated values of the explanatory variables at time $s$, as a consequence of the choice made in the current time interval $t$. This anticipation is represented by a probability density function (pdf):

$$f_{\tilde{X}_{ns}(t)}(x|y_{nt}, \tilde{x}_{nt}),\ t < s \leqslant T, \tag{5}$$

where the notation $\tilde{X}_{ns}(t)$ emphasizes that the anticipation of the values of the variables at time $s$ may change over time $t$. This reflects the ability of the individual to update their expectations of the future variables as time unfolds. Note that the decision maker has many possible ways to anticipate the future. For example, they may consider hypothetical scenarios based on prior experiences or perceptions, or consult online databases/ information services, etc. These sources of information can be included in $\tilde{x}_{nt}$ without loss of generality.

Assuming the individual $n$ is rational, she evaluates a vector of utility (aka payoff or reward)

$$\tilde{u}_{nt} = \tilde{u}(\tilde{x}_{nt}) \in \mathbb{R}^J, \tag{6}$$

[2]Note that if $\mathcal{C}$ contains alternatives that are not available to all individuals at all times, then $\mathcal{T}$ also contains trajectories that cannot be chosen.

for each time interval. The function $\tilde{u}$ captures the decision maker's individual preferences, including how the variables in $\tilde{x}_{nt}$ affect her derived utility of each alternative. If $t$ is the current time interval, $\tilde{u}(\tilde{X}_{ns}(t))$ is the individual's anticipated future utility at time $s > t$, based on her anticipation of the values of variables considered at time $t$. The lower case notation for $\tilde{u}$ emphasizes that the only source of randomness in the utility comes from the anticipated values of the variables.

In the final time interval $t = T$, the decision maker simply maximizes the utility at time $T$

$$\max_{y_T \in \mathcal{Y}} y_T^T \tilde{u}(\tilde{x}_{nT}), \tag{7}$$

where $y_T$ represents a fixed choice at time $T$. For all remaining time intervals $t < T$, the decision maker maximizes the total expected (discounted) utility, that is the utility at time $t$, plus the expected utility in future time intervals,

$$\max_{y \in \mathcal{T}_t} y_t^T \tilde{u}(\tilde{x}_{nt}) + E_{\tilde{X}_{n,t+1:T}(t)} \left( \sum_{s=t+1}^{T} \rho_n^{s-t} y_s^T \tilde{u}(\tilde{X}_{ns}(t)) \right), \; t < T, \tag{8}$$

where: (i) $\mathcal{T}_t$ is the set of all trajectories starting at time $t$ as defined by (4), (ii) $y = (y_t, y_{t+1}, ..., y_T)$ represents a single possible trajectory with a fixed choice $y_t$ and anticipated future choices $y_s$, and (iii) $0 \leqslant \rho_n \leqslant 1$ is a discount factor. Note that the decision maker does not commit to the anticipated choices $y_s$ where $s > t$, as they are based on anticipated information. The trajectory $y_n = y_{n,0:T}$ chosen by the individual is hence the result of solving (8) at each time $t = 1, \ldots, T$. The diagram in Figure 1 illustrates the point of view of the decision maker, including the relationships between the variables introduced in this section.

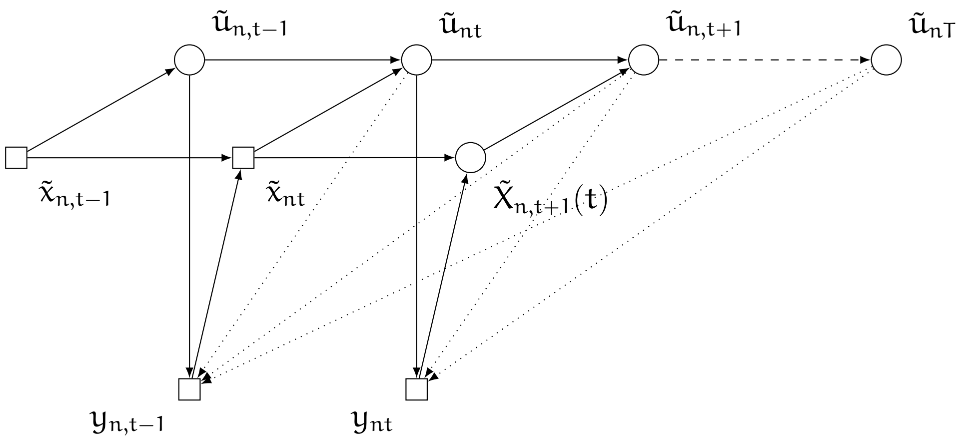


Figure 1: Illustration of the variables involved in the decision making process at the current time interval $t$. The shape of a node represents its nature: Circle = latent variable, square = observed variable. For clarity of the diagram, the dependence of the choice on future (anticipated) values of the utility are shown as dotted arrows.

The value of the discount factor $\rho_n$ can reflect different types of behaviour. A value of $\rho_n = 0$ describes a fully myopic behaviour, where the individual evaluates only the utility at time $t$ without taking into account the future consequences of their decisions.

Conversely, a value of $\rho_n = 1$ implies a fully forward-looking behaviour, where the individual values equally the utility at time $t$ and the expected utility in future time intervals. Values $0 < \rho_n < 1$ represents limited forward-looking planning, where the individual accounts for the expected utility in future time intervals, but places decreasing importance on the expected utility as $s$ (and therefore the prediction horizon $s - t$) increases.

In the next section we take the point of view of the analyst. We show that the decision-making problem (8) can be formulated as a dynamic programming problem using Bellman's principle of optimality. Note that, for that purpose, the additive specification (8) of the utility function of the trajectory is critical.

# 3 The point of view of the analyst

The objective of the analyst is to specify a model that can accurately predict individuals' sequences of unobserved choices. In addition, it is often desirable that the models and the resulting predictions are interpretable. Crucially, the prediction and estimation problems must be computationally tractable. In this section we describe the individual's choice problem introduced in the previous section, acknowledging that the analyst does not have perfect knowledge of its elements. Furthermore, for computational tractability, we introduce a dynamic programming formulation of the optimization problem (8). Section 4 is devoted to a parametric formulation and maximum likelihood estimation.

Based on Bellman's principle of optimality (Bellman, 1952), the idea of dynamic programming is to construct the optimal trajectory for (8) piece by piece (see, e.g., Bertsekas, 2017). More precisely, this is achieved by solving a backward recursive formulation of *value functions* defined by the Bellman equation.

In this recursive context, we specify the model for the anticipation of the future variables (5) one time interval at a time. For $s \geqslant t$, the random variable $\tilde{X}_{n,s+1}(t)$ represents the anticipation of the explanatory variables for time interval $s + 1$, performed at time $t$. It is characterized by the pdf

$$f_{\tilde{X}_{n,s+1}(t)}(x|y_{ns}, \tilde{X}_{ns}(t)), \; t \leqslant s < T. \tag{9}$$

This is the analyst's attempt to approximate (5) in a recursive way. As in (5), the notation $\tilde{X}_{n,s+1}(t)$ emphasizes that the anticipation of the values of the variables at time $s + 1$ may change over time $t$. Note that, if $s = t$, the explanatory variables $\tilde{x}_{nt}$ are observed and not anticipated, and

$$\tilde{X}_{nt}(t) = \tilde{x}_{nt}. \tag{10}$$

The analyst does not have access to the true utility functions $\tilde{u}$. Furthermore, they do not have access to all of the variables considered by the decision-maker, and instead only have access to a vector of observable variables, which we denote $x_{nt}$ (see Section 4.3). As such, the values of the utility are modeled using random variables denoted $U$. The randomness of the utility functions is motivated by random utility theory (e.g., Manski, 1973, Manski, 1977).

If the present time is $t$, we recursively define, for a considered interval $s \geqslant t$, a *global utility* $U'_i(\tilde{X}_{ns}(t))$ for each alternative $i$, which involves an *instantaneous* utility $U_i(\tilde{X}_{ns}(t))$, and a *future* utility $W_i(\tilde{X}_{ns}(t))$. This formulation reflects the fact that the anticipation of the future variables in (5) and (9) is not constant over time. At the present time $t$, the individual can consider the choice at present or future time interval $s$, where $t \leqslant s \leqslant T$. The anticipated values of the variables at time $s$ are represented by $\tilde{X}_{ns}(t)$. The instantaneous utility $U_i$ represents the utility the individual expects to derive from choosing alternative $i \in \mathcal{C}$, based on the anticipated values of the variables at time $s$. The future utility $W_i$ then represents the expected maximum total utility over subsequent time intervals $s+1, ..., T$, assuming alternative $i$ is chosen at time $s$, where $s < T$. To reflect this, we also define a *value function* $w^*_{ns}(t)$, that captures the expected maximum global utility of choosing alternative $i$ at time $s$, considered at the current time $t$. A recursive definition is the key to unifying the notation of the general dynamic choice model, and allows us to derive the choice probabilities using Bellman's equation.

The recursive definition works backwards. At time interval $s = T$, the anticipated utility is simply the instantaneous utility, that is

$$U'_i(\tilde{X}_{nT}(t)) = U_i(\tilde{X}_{nT}(t)), \; t \leqslant T. \tag{11}$$

The value function is defined as the expected optimal value of the problem (8) solved by the decision maker:

$$w^*_{nT}(t) = w^*(\tilde{X}_{nT}(t)) = E_U[\max_{j \in \mathcal{C}} U'_j(\tilde{X}_{nT}(t))], \; t \leqslant T. \tag{12}$$

For $t \leqslant s < T$, the global utility of alternative $i$ is defined as

$$U'_i(\tilde{X}_{ns}(t)) = U_i(\tilde{X}_{ns}(t)) + \rho_n W_i(\tilde{X}_{ns}(t)), \; t \leqslant s < T, \tag{13}$$

where: (i) $U_i(\tilde{X}_{ns}(t))$ is the instantaneous utility for time $s$, as evaluated at time $t$, (ii) $W_i(\tilde{X}_{ns}(t))$ is the utility to be obtained in the future if alternative $i$ is chosen at time $s$, and (iii) $\rho_n$ is the individual discount factor introduced in (8). Furthermore, the expected future utility when choosing $i$ at $s$ is

$$\begin{aligned} W_i(\tilde{X}_{ns}(t)) &= E_{\tilde{X}_{n,s+1}(t)}[w^*_{n,s+1}(t) | y_{ns} = \delta_i], \\ &= \int_x w^*_{n,s+1}(x) f_{\tilde{X}_{n,s+1}(t)}(x | \delta_i, \tilde{X}_{ns}(t)) dx, \; t \leqslant s < T, \end{aligned} \tag{14}$$

where $f_{\tilde{X}_{n,s+1}(t)}$ is the pdf (9) of $\tilde{X}_{n,s+1}(t)$. Then, the value function at time $s$ is defined as

$$w^*_{ns}(t) = w^*(\tilde{X}_{ns}(t)) = E[\max_{j \in \mathcal{C}} U'_j(\tilde{X}_{ns}(t))], \; t \leqslant s < T. \tag{15}$$

By substituting (11) into (12) and (13) into (15) we obtain the full form value function at time $t$:

$$w^*_{ns}(t) = \begin{cases} E[\max\limits_{j \in \mathcal{C}} U_j(\tilde{X}_{ns}(t)) + \\ \qquad \rho_n \int_x w^*_{n,s+1}(x) f_{\tilde{X}_{n,s+1}(t)}(x | \delta_j, \tilde{X}_{ns}(t)) dx], & s < T, \\ E[\max\limits_{j \in \mathcal{C}} U_j(\tilde{X}_{nT}(t))], & s = T. \end{cases} \tag{16}$$

The choice model, that is, the probability of choosing $i_{nt}$ at time $t$ is

$$P(i_{nt} | \tilde{x}_{nt}, \mathcal{C}) = \text{Prob}(U'_i(\tilde{x}_{nt}) \geqslant U'_j(\tilde{x}_{nt}), \forall j \in \mathcal{C}), \tag{17}$$

or, equivalently,

$$P(y_{nt}|\tilde{x}_{nt},\mathcal{C}) = \text{Prob}(y_{nt}^T U'(\tilde{x}_{nt}) \geqslant \delta_j^T U'_{nt}(\tilde{x}_{nt}), \forall j \in \mathcal{C}). \tag{18}$$

Thus far, we have not made any distributional assumptions on the random variables $\tilde{X}_{n,s+1}(t)$ and $U_i(\tilde{X}_{ns}(t))$. In the next section, we propose parametric model specifications and discuss maximum likelihood estimation.

# 4 A general parametric model and estimation

We introduce in this section the modelling assumptions that allow the analyst to derive a likelihood function associated with the data.

## 4.1 Parametric model

The first assumption of the parametric model is that the distribution (9) of the future explanatory variables can be modeled with a Markov chain

$$\tilde{X}_{n,s+1}(t) = h(y_{ns}, \tilde{X}_{ns}(t); \theta_h) + \alpha_n^x + \lambda_\nu^{s+1-t}\nu_{n,s+1}, \ t \leqslant s < T, \tag{19}$$

where: (i) $\theta_h$ and $\lambda_\nu \geqslant 1$ are parameters, (ii) $\tilde{X}_{nt}(t) = \tilde{x}_{nt}$, (iii) $\alpha_n^x$ are i.i.d. across $n$ with pdf $f_{\alpha^x}(x;\theta_{\alpha^x})$, and (iv) $\nu_{n,s+1}$ are i.i.d. across $n$ and $s$, with pdf $f_\nu(x;\theta_\nu)$, and independent from $t$. The first term captures the dynamics of the incremental anticipation, independently of $s-t$. For instance, it may include the impact of the purchase of item $i$ on the income available for the next time interval. The second term is an error term specific to individual $n$ and constant over time[3]. The third term is an error term defined such that its variance increases with $s$ being further away in the future[3] (i.e., as $s-t$ increases), capturing the fact that the quality of the anticipation decreases with time. Note that the presence of $\alpha_n^x$ explicitly captures serial correlation of the error terms, so that the assumption that $\nu_{ns}$ are independent across $s$ is acceptable[4].

Recall from (5) that $\tilde{X}_{ns}$ can include both historical values of the attributes of the alternatives, as well as sources of information to consult the future, such as online databases. By including historical values in $\tilde{X}_{ns}$, the Markov chain in (19) can represent any possible form of the anticipation of the future explanatory variables.

If $f_\nu$ is the pdf of $\nu_{n,s+1}$, we have

$$f_{\tilde{X}_{n,s+1}(t)}(x|y_{ns},\tilde{X}_{ns}(t),\alpha_n^x) = \frac{1}{\lambda_\nu^{s+1-t}} f_\nu\left(\frac{x - h(y_{ns},\tilde{X}_{ns}(t);\theta_h) - \alpha_n^x}{\lambda_\nu^{s+1-t}};\theta_\nu\right). \tag{20}$$

Note that $f_\nu$ is not indexed by $s$, $n$ or $t$, because of the i.i.d. assumption. The variations across time and individuals are explicitly captured by the specification (19).

[3]Note that the two superscripts in (19) have different meanings. The superscript $x$ in $\alpha_n^x$ indicates that this error term relates to the explanatory variables (as opposed to $\alpha_{in}^U$ introduced in (21) which relates to the utilities). Conversely, the $s+1-t$ in $\lambda_\nu^{s+1-t}$ is an exponent which increases the variance of $\nu_{n,s+1}$ as $s-t$ increases.

[4]Note that in the presence of very long observation periods, the random parameters in (19) can be replaced by fixed parameters (see Section 4.2).

For the utility function, it is convenient to capture the sources of randomness using an additive specification. We model (13) as

$$U'_i(\tilde{X}_{ns}(t)) = V'_i(\tilde{X}_{ns}(t)) + \alpha^U_{in} + \lambda_\varepsilon^{s-t}\varepsilon_{ins},\ t \leqslant s \leqslant T \tag{21}$$

where the first term is deterministic, conditional on $\tilde{X}_{ns}(t)$. It is defined as

$$V'_i(\tilde{X}_{ns}(t)) = V_i(\tilde{X}_{ns}(t)) + \rho_n W_i(\tilde{X}_{ns}(t)),\ t \leqslant s \leqslant T. \tag{22}$$

The error term has two components: $\alpha^U_{in}$, i.i.d. across $n$ and constant over $t$, with pdf $f_{\alpha^U}(x;\theta_{\alpha^U})$, and $\varepsilon_{ins}$, i.i.d. across $n$ and $s$, with pdf $f_\varepsilon(x;\theta_\epsilon)$, and independent from $t$. Similarly to the specification of the future variables (19), the term $\alpha^U_{in}$ captures serial correlation[5], and it is explicitly assumed that the variance of the error term increases by a factor $\lambda_\varepsilon$ at each time interval[6]. .

The type of choice model is implied by the assumption on the error terms $\varepsilon_{nt}$. For example, if they are assumed to be i.i.d. Extreme Value distributed with scale parameter $\mu$, then the value function (15) is

$$\begin{aligned} w^*(\tilde{X}_{ns}(t)) &= E_{\alpha^U_n}[E_{\varepsilon_{ns}}[\max_{i\in\mathcal{C}} U'_{ins}(\tilde{X}_{ns}(t))]] \\ &= E_{\alpha^U_n}\left[\frac{1}{\mu_{st}}\ln\sum_{i\in\mathcal{C}}\exp(\mu_{st}(V'_{int}(\tilde{X}_{ns}(t)) + \alpha^U_{in}))\right], \end{aligned} \tag{23}$$

where

$$\mu_{st} = \frac{\mu}{\lambda_\varepsilon^{s-t}} \tag{24}$$

is the scale parameter. In this case the choice model (18) is a mixture of logit models,

$$P(i_{nt}|\tilde{x}_{nt}) = E_{\alpha^U_n}\left[P(i_{nt}|\tilde{x}_{nt},\alpha^U_n)\right], \tag{25}$$

where

$$P(i_{nt}|\tilde{x}_{nt},\alpha^U_n) = \frac{\exp(\mu_{st}(V'_{int}(\tilde{x}_{nt}) + \alpha^U_{in}))}{\sum_{j\in\mathcal{C}}\exp(\mu_{st}(V'_{jnt}(\tilde{x}_{nt}) + \alpha^U_{jn}))}. \tag{26}$$

It is also assumed that the error components $\alpha^x_n$, $\nu_{n,s+1}$, $\alpha^U_{in}$, and $\varepsilon_{in}$ are all independent from each other.

The unknown parameters are:

- the parameters of the utility functions, that have not yet been introduced, and that we denote by $\beta$,
- the discounting parameters $\rho_n$,
- the parameters of the variables anticipation model $\theta_h$,
- the variance inflation parameters $\lambda_\nu$ and $\lambda_\varepsilon$,

[5] As with (19), in the presence of very long observation periods, the random parameters in (21) can be replaced by fixed parameters (see Section 4.2).

[6] As with (19), the superscripts in (21) have different meanings: The superscript $U$ in $\alpha^U_{in}$ indicates that this error term relates to the utilities, whilst the $s-t$ in $\lambda_\varepsilon^{s-t}$ is an exponent which increases the variance of $\epsilon_{ins}$ as $s-t$ increases.

- the parameters of the distribution of the individual effects (these are commonly referred to as agent effects), $\theta_{\alpha^x}$ and $\theta_{\alpha^u}$, and
- the parameters of the pdf of $\nu$, $\theta_\nu$.

We denote by $\theta$ the vector of all these parameters.

## 4.2 Agent effects

In the previous section, we introduced the agent effects $\alpha_n^u$ and $\alpha_n^x$ as random variables, which are distributed across the population. This is known as a *random-effects* model. For example, the utility agent effects could be distributed according to a normal distribution

$$\alpha_n^u \sim N(0, \Sigma). \tag{27}$$

However, in the presence of very long observation periods where there are many observations per individual, the agent effects can instead be modelled as fixed. This is known as a *fixed-effects* model. Models with fixed effects consider $\alpha_n$ as individual-specific vectors of unknown parameters to be estimated from data. For example, in a fixed-effects model, $\alpha_n^u$ would contain one parameter for each alternative $i \in J$, with one parameter normalised to zero, such that for $N$ individuals or homogeneous classes of individuals, $N(J-1)$ parameters must be calculated.

Whilst dynamic models typically make use of random effects (see Section 6), there has been extensive investigation of the use of fixed effects for static models with panel data, where the sequence of choices made by an individual is considered independent over time. Static models can be considered as a restricted version of the general parametric model, where the individual behaves myopically (i.e. $\rho = 0$) and the choice made is independent of previous time periods. For example, consider a stated preference survey, where each respondent provides their indicated choice for a sequence of independent hypothetical choice situations. Under certain conditions (assuming in particular that the effect of survey fatigue is low), each individual's choice process across the hypothetical choice situations could be assumed to be constant. However, the inter-individual heterogeneity could be significant, and the modeller may wish to estimate the agent effects. These scenarios can be investigated using static choice models with either fixed or random agent effects. For a more detailed overview of the use of fixed and random effects in static choice models, we direct the reader to Greene (2001).

## 4.3 Maximum likelihood estimation

It is assumed that the analyst has access to *panel data* or *longitudinal data* for a sample of $N$ individuals in the population. The observation period starts at time $t_b$ and ends at time $t_e$, such that $0 \leqslant t_b < t_e \leqslant T$. The number of time intervals in the sample is hence $T_s = t_e - t_b + 1$. Note that if $t_b = t_e$, data would be available only for one time interval. In that case, it would be called *cross-sectional* data that do not provide information about the time dimension. For simplicity of the notation, we assume that the panel data is *complete*, in the sense that data is available for all time intervals between $t_b$ and $t_e$, and *balanced*, meaning that all the explanatory variables for all individuals are available at each time interval during the observation period. However, this is not a strict requirement for the estimation of dynamic choice models.

The analyst uses the panel data to estimate the parameters of the model. For each time interval $t$ during the observation period, and for each individual $n$, the analyst has access to the observed choice, represented by the binary vector $y_{nt}$ and a vector of *observed* explanatory variables $x_{nt}$. As with $y_n$, we use the notation $x_n = x_{n,0:T}$ to denote sequence of explanatory variables for individual $n$ over time. Note, unlike $\tilde{x}_{nt}$, $x_{nt}$ can only include variables observable to the analyst, and must be truly exogenous from the model. As such, $x_{nt}$ is considered separately and distinctly from historic values of observed choices $y_{nt}$ and utilities $U_{nt}$.

In order to estimate the parameters from data by maximum likelihood, we derive the contribution to the likelihood of the observations related to individual $n$, $\ln \text{Prob}(y_{n,t_b:t_e}, x_{n,t_b:t_e}|\theta)$. We isolate the agent effects, that are constant over time, so that the contribution of individual $n$ to the conditional likelihood function is

$$\ell_n(\theta) = \ln \text{E}_{\alpha^x,\alpha^U}[\text{Prob}(y_{n,t_b:t_e}, x_{n,t_b:t_e}|\alpha^x, \alpha^U, \theta)]. \tag{28}$$

We then exploit the recursive definition of the model, and the assumptions of independence of the error components over time, such that

$$\begin{aligned}
&\text{Prob}(y_{n,t_b:t_e}, x_{n,t_b:t_e}|\alpha^x, \alpha^U, \theta) = \\
&\quad \text{Prob}(y_{n,t_b}, x_{n,t_b}|\alpha^x, \alpha^U, \theta) \prod_{t=t_b+1}^{t_e} \text{Prob}(y_{nt}, x_{nt}|y_{n,t_b:t-1}, x_{n,t_b:t-1}, \alpha^x, \alpha^U, \theta),
\end{aligned} \tag{29}$$

where $y_{n,t_b:t-1}$ and $x_{n,t_b:t-1}$ represent the entire history of choices and explanatory variables respectively from time $t_b$ to time $t-1$. Note that the first observation (at $t = t_b$) is the *initial condition*, which cannot be conditioned on previous data, and so is included separately in (29). This is discussed in more detail in Section 5.1.

The joint probability in (29) can be expressed as the product of the marginal probability from the anticipation of the explanatory variables and a conditional choice probability using Bayes theorem

$$\begin{aligned}
&\text{Prob}(y_{nt}, x_{nt}|y_{n,t_b:t-1}, x_{n,t_b:t-1}, \alpha^x, \alpha^U, \theta) = \\
&\qquad \text{Prob}(x_{nt}|y_{n,t_b:t-1}, x_{n,t_b:t-1}, \alpha^x, \alpha^U, \theta) \times \\
&\qquad\qquad \text{Prob}(y_{nt}|x_{nt}, y_{n,t_b:t-1}, x_{n,t_b:t-1}, \alpha^x, \alpha^U, \theta).
\end{aligned} \tag{30}$$

It is not feasible, neither from a computational nor a data perspective, to estimate models conditional on the full history of explanatory variables/decisions. In Section 4.1, we introduced the Markov chain for the anticipation of the explanatory variables, which models the anticipation based on only the previous time period. Substituting $t = s$ (so that the considered time is the current period) followed by $s = t - 1$ (to shift one time period back) into (20) gives

$$\begin{aligned}
\text{Prob}(x_{nt}|&y_{n,t_b:t-1}, x_{n,t_b:t-1}, \alpha^x, \alpha^u, \theta) \\
&= P(x_{nt}|y_{n,t-1}, x_{n,t-1}, \alpha^x, \theta) \\
&= \frac{1}{\lambda_\nu} f_\nu \left( \frac{x_{nt} - h(y_{n,t-1}, x_{n,t-1}; \theta_h) - \alpha^x_n}{\lambda_\nu}; \theta_\nu \right).
\end{aligned} \tag{31}$$

The choice model can be similarly simplified. For example, in the next section we introduce a choice model which depends only on the current values of $x_{nt}$ as well as the choice from the previous time interval, so that

$$\text{Prob}(y_{nt}|x_{nt}, y_{n,t_b:t-1}, x_{n,t_b:t-1}, \alpha^x, \alpha^u, \theta) = P(y_{nt}|y_{n,t-1}, x_{nt}, \alpha^u, \beta, \lambda_\varepsilon, \rho). \quad (32)$$

Models with $\rho_n > 0$ are particularly challenging to estimate because the choice probabilities depend on the recursively defined expected future utilities. The solutions to the expected future utilities hence need to be computed when evaluating the likelihood function. Rust (1987) propose the Nested Fixed Point Estimator (NXFP) that is based on an outer and an inner algorithm. The former searches over the parameter space maximizing the likelihood function while the latter solves the value functions. The estimation problem is hence computationally costly. With the objective to reduce the computational burden, several alternatives to the NXFP algorithm have been proposed in the literature (e.g., Hotz and Miller, 1993; Hotz et al., 1994; Imai et al., 2009; Keane and Wolpin, 1994; Su and Judd, 2012).

# 5 Habitual behaviour and learning

Panel data provide information about the evolution of choice behaviour over time, and so present the opportunity to capture the development of learning and the role of habits. Learning and habits determine how past experiences impact an individual's decisions. Capturing learning and habits within a model therefore requires past experiences to be included in the utility function (21). This presents two key questions:

1. What variables can we use to capture past experiences?

2. How far in the past should we consider?

For the first question, there are many variables that could be used to capture past experience, including previous choices, explanatory variables, and latent variables or states. We consider here two possibilities: the previous choices made and previous values of the utility. For the second question, as discussed in the previous section, an individual's decision at time $t$ could be dependent on all of their past experiences from periods $0, ..., t-1$. However, in order to enable a recursive model definition that can be used to predict choice sequences of arbitrary length, we must instead consider the past experiences from a fixed number $k$ of lagged time intervals. A higher value of $k$ represents a more flexible model. However, to estimate a model with $k$ lagged time intervals, the first $k$ observations for each individual in the data must be assumed as given, and so are not available for model estimation. This therefore effectively reduces the available data for model estimation. As such, it is typical, with dynamic choice models to apply the *Markov assumption* by fixing $k = 1$. It means that, at time interval $t$, the entire past is modeled using only the previous time interval $t - 1$. There are, however, examples in the literature where values of $k > 1$ are used (see Sections 6.2 and 6.3).

When we combine the Markov assumption with the use of the choice to define past experience, we obtain the *Markov model*. When we combine it instead with the use of the latent utility to define past experience, we obtain the *hidden Markov model*. We first present the Markov model, alongside a related econometric issue called the *initial condition problem*. We then present the hidden Markov model and introduce a solution algorithm called *particle filtering*.

## 5.1 The Markov model and the initial condition problem

For the Markov model, we explicitly include the previous choice as an explanatory variable to the utility function. As in 4, we set $s = t$ to consider only decisions made in the current time period. Equation (21) therefore becomes

$$U'_i(\tilde{x}_{nt}) = V'_i(\tilde{x}_{nt}) + \eta_i y_{n,t-1} + \alpha^U_{in} + \varepsilon_{int}, \; t \leqslant T. \tag{33}$$

Note the parallel between the Markov chains in (33) and (19). The Markov chain in (33) assumes that the utility in the current time interval is dependent on the choice in the previous time interval, in order to model habitual behaviour and learning. Meanwhile, the Markov chain in (19) assumes that the anticipated values of the explanatory variables in the next time interval are dependent on their values (as well as the choice made) in the current time interval, in order to model forward-planning behaviour.

A major difficulty in modelling the dynamics of choice, and the influence of the past on current decisions, arises when the observation period does not include the entire history of the process. In particular, everything that happened between time 0 and time $t_b$ is captured only by the observation of the choice at time $t_b$. This may lead to erroneous interpretation of the choice.

To demonstrate this, consider two individuals with strong habits, so that their choice made today is largely explained by their choice made yesterday. For instance, out of two commuters, one might be a "car lover" and another one a "public transportation lover". These commuters would stick to their preferred mode except in rare circumstances, even if that mode is slower or more expensive than the alternatives. If the observation period does not include the day when each commuter made their choice for the first time, the analyst would not have access to the variables explaining that choice. It may therefore appear that these individuals prefer slower or more expensive alternatives. In turn, this would impact the estimated coefficients of the model variables. The unobserved variables explaining the first choice, which explain the differences in taste, actually belong to the agent effects $\alpha^x_n$ and $\alpha^U_{in}$. For instance, the "car lover" has a large $\alpha^U_{in}$ for the car alternative, while the "public transportation lover" has a large $\alpha^U_{in}$ for the public transportation alternative. Consequently, the analyst cannot assume the same distribution for all individuals in the population. Doing so would cause an *endogeneity* issue, as the random term $\alpha^U_{in}$ would be correlated with the initial choice $y_{nt_b}$. This is called the *initial condition problem.*

Wooldridge (2005) proposes to model $\alpha^x_n$ and $\alpha^U_{in}$, conditional on $y_{nt_b}$. For instance, $\alpha^x_n$ can be represented as

$$\alpha^x_n = a_x y_{nt_b} + b^T_x x'_n + \zeta^x_n, \tag{34}$$

where: (i) $x'_n$ are observed socio-economic characteristics of individual $n$; (ii) $\zeta_n$ is assumed to be normally distributed and independent from $y_{nt_b}$:

$$\zeta^x_n \sim N(0, \Sigma_\zeta); \tag{35}$$

and (iii) $a_x$ and $b_x$ are unknown parameters to estimate from data. The agent effect $\alpha^U_{in}$ can be modeled in a similar way:

$$\alpha^U_{in} = a_U y_{nt_b} + b_U^T x'_n + \zeta^U_n. \tag{36}$$

This specification addresses the initial condition problem presented here, caused by serial correlation.

Although this issue is commonly considered in the analysis of panel data, it is actually a more general issue that applies to all models capturing learning, due to the impossibility to observe the whole history of experiences (Guevara et al., 2018). It also applies to contexts where each individual is associated with multiple observations, such as stated preference surveys.

## 5.2 The hidden Markov model and particle filtering

A second way to model the evolution of learning and habitual behaviours consists in directly adding the previous utility function (i.e. a continuous latent variable) as an explanatory variable to the utility function. More specifically, for $s = t$, the utility function (33) becomes

$$U'_i(\tilde{x}_{nt}) = \tilde{U}_{int} + \varepsilon_{int}, \tag{37}$$

where: (i) the evolution of the vector of utilities $\tilde{U}_{nt}$ over time is modeled as

$$\tilde{U}_{int} = V'_i(\tilde{x}_{nt}) + \gamma_i \tilde{U}_{n,t-1} + \eta_i y_{n,t-1} + \alpha^U_{in} + \xi_{nt}; \tag{38}$$

(ii) the random vectors $\xi_{nt}$ are i.i.d. normal, that is, for all $n$ and $t$

$$\xi_{nt} = \Sigma_\xi \omega; \tag{39}$$

(iii) $\Sigma_\xi$ is the Cholesky factor of the variance covariance matrix; and (iv) $\omega \sim N(0, I)$ follows a standard normal distribution. Note that the recursive definition of the vector of utilities in (38) involves the whole sequence of previous utility functions.

This model is called a *hidden Markov model*. It is a Markov model where some state variables are latent, i.e., not observed. In our context, the latent state variables are the utility functions.

Note that this formulation significantly complicates the calculation of (29). Indeed, the choice model in (30) now involves the full trajectory of utility functions:

$$P(y_{nt}|y_{n,t-1}, x_{nt}, \alpha^U_n, \beta, \lambda_\varepsilon, \rho) = E_{\tilde{U}_{n,t_b:t-1}}[P(y_{nt}|\tilde{U}_{n,t_b:t-1}, x_{nt}, \alpha^U_n, \beta, \lambda_\varepsilon, \rho)]. \tag{40}$$

The calculation of the expectation involves a multifold integral with $t - t_b$ dimensions, which is in general too complicated to handle. In order to simplify it, we again need a recursive definition of the model.

We present a method called *particle filtering*, inspired by the work of Kalman (1960), that is designed to update the estimates at each time interval. We write (38)–(39) as

$$\omega(\tilde{U}_{nt}, \tilde{U}_{n,t-1}) = \Sigma_\xi^{-1}(\tilde{U}_{nt} - V'(\tilde{x}_{nt}) - \gamma \tilde{U}_{n,t-1} - \eta y_{n,t-1} - \alpha^U_n). \tag{41}$$

Using this change of variables, we can write the density of $\tilde{U}_{nt}$ conditional on $\tilde{U}_{n,t-1}$, $y_{n,t-1}$, $\alpha_n^U$ and $\tilde{x}_{nt}$ as

$$f_{\tilde{U}_{nt}}(u|\tilde{U}_{n,t-1}, y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt}) = \frac{1}{|\Sigma_\xi|}\phi(\omega(u, \tilde{U}_{n,t-1})), \tag{42}$$

where: (i) $\omega$ is defined by (41), (ii) $|\Sigma_\xi|$ is the determinant of the matrix $\Sigma_\xi$, and (iii) $\phi(\cdot)$ is the pdf of the standard normal distribution:

$$\phi(x) = (2\pi)^{\frac{J}{2}} e^{-\frac{1}{2}x^T x}. \tag{43}$$

If we integrate out $\tilde{U}_{n,t-1}$, we obtain

$$f_{\tilde{U}_{nt}}(u|y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt}) = \frac{1}{|\Sigma_\xi|}\int_v \phi(\omega(u; v)) f_{\tilde{U}_{n,t-1}}(v|y_{n,t-1}, y_{n,t-2}, \alpha_n^U, \tilde{x}_{n,t-1})dv, \tag{44}$$

where $\omega(u; v)$ is defined by (41) and the distribution of $\tilde{U}_{n,t-1}$, conditional on the choices of the two previous time intervals, is defined below. In the particle filtering literature (Julier & Uhlmann, 1997), (44) is called *state prediction*. In our context, the (latent) state is the utility.

The choice model (40) is now written as a mixture model:

$$P(y_{nt}|y_{n,t-1}, x_{nt}, \alpha_n^U, \beta, \lambda_\varepsilon, \rho) = \int_u P(y_{nt}|u, y_{n,t-1}, x_{nt}, \alpha_n^U, \beta, \lambda_\varepsilon, \rho) f_{\tilde{U}_{nt}}(u|y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt})du, \tag{45}$$

where:

$$\text{(i) } P(y_{nt}|u, y_{n,t-1}, x_{nt}, \alpha_n^U, \beta, \lambda_\varepsilon, \rho) = \frac{\exp(\mu_{st}u_i)}{\sum_{j\in\mathcal{C}} \exp(\mu_{st}u_j)} \tag{46}$$

is the logit model (26) expressed as a function of $u$, which is a realization of $\tilde{U}_{nt}$; and (ii) $f_{\tilde{U}_{nt}}(u|y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt})$ is defined by the state prediction (44).

In order to propagate the filter to the next time interval, we need

$$f_{\tilde{U}_{nt}}(u|y_{nt}, y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt}) \tag{47}$$

to apply the state prediction (44). It can be obtained by Bayes' theorem:

$$f_{\tilde{U}_{nt}}(u|y_{nt}, y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt}) = \frac{\text{Prob}(y_{nt}|u, y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt}) f_{\tilde{U}_{nt}}(u|y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt})}{\text{Prob}(y_{nt}|y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt})}, \tag{48}$$

where the involved quantities are the conditional choice probability (46) and the state prediction (44) at the numerator, and the choice probability (45) at the denominator.

The particle filtering is initialized with the distribution of the utility function of the first time internal:

$$f_{\tilde{U}_{nt_b}}(u|\alpha_n^U, \tilde{x}_{nt_b}). \tag{49}$$

For each time interval $t = t_b + 1, \ldots, t_e$, the procedure is as follows:

1. We have access to the density of the utility of the previous time interval

$$f_{\tilde{U}_{n,t-1}}(u|y_{n,t-1}, \alpha_n^U, \tilde{x}_{n,t-1}). \tag{50}$$

2. We use the state prediction (44) to calculate

$$f_{\tilde{U}_{nt}}(u|y_{n,t-1}, \alpha_n^U, \tilde{x}_{nt}). \tag{51}$$

3. We calculate the mixture of logit models (45) to obtain the contribution of time interval $t$ to the likelihood.

4. We prepare the density of the utility for the next time interval using (48) to obtain

$$f_{\tilde{U}_{nt}}(u|y_{nt}, \alpha_n^U, \tilde{x}_{nt}). \tag{52}$$

The example discussed here includes the previous value of the utility in the utility function, and shows how particle filtering can be used to address the complexity of estimating the model. Particle filtering can be used for any model where a *latent variable* that changes over time is included in the utility function. This includes other continuous latent variables, such as the agent effects $\alpha_n^U$, or transitions between discrete latent states or classes. We present examples from the literature of both in Section 6.3.

# 6 Links to existing models

The general parametric model introduced in this chapter can be used to derive different types of dynamic choice models. We start by introducing examples of forward looking models, followed by Markov and hidden Markov models. In each case, we present how different assumptions made on the parameters of the general model can be used to derive different example models from the literature. We then summarise how these models have been applied in selected relevant studies, and discuss which applications, data, and choice situations each model is appropriate for.

## 6.1 Forward looking models

The first notable example of a forward looking dynamic choice model estimated in the literature is that of Rust (1987), who investigates the sequential choices of a single decision maker for bus engine replacement timing. Here the decision variable $y_{nt}$ is a binary variable that represents the decision to replace the engine for bus $n$ in month $t$ or not, and the only explanatory variable is the mileage $x_{nt}$ since last engine replacement of bus $n$ at month $t$. Rust's model can be obtained by making the following assumptions on (21):

1. the error terms $\varepsilon_{nt}$ are i.i.d. Extreme Value to give the logit model (as exemplified in (23)), with constant variance, so that $\lambda_\varepsilon = 1$;

2. there is no serial correlation of the utilities to ensure additive separability, i.e., $\alpha_n^U = 0$; and

3. there is no serial correlation of the anticipation of the future variables to ensure conditional independence, so that $\alpha_n^x = 0$ and $\lambda_\nu = 0$.

This gives the following form of the global utility:

$$U_i'(\tilde{X}_{ns}(t)) = V_i(\tilde{X}_{ns}(t)) + \rho_n W_i(\tilde{X}_{ns}(t)) + \varepsilon_{ins},\ t \leqslant s \leqslant T. \tag{53}$$

For the specific example in the paper, the deterministic portion of the instantaneous utility of replacing the engine is given by

$$V(\tilde{X}_{ns}(t)) = \beta_0 + g(x_t, \beta_x) \tag{54}$$

where multiple different functional forms (linear, quadratic, cubic, square root, power, hyperbolic, mixed, and non-parametric) are tested for $g(x_t, \beta_x)$. Note that as the decision variable is binary, the $i$ subscript can be dropped as we only need to calculate the utility for making the engine replacement.[7] The anticipation of the mileage since last replacement at month $s + 1$ given mileage since last replacement at month $s$ is then defined as

$$f_{\tilde{X}_{n,s+1}(t)}(x|y_{ns}, \tilde{X}_{ns}(t)) = \theta e^{\theta(x_{s+1}-(1-y_s)x_s)}\ t \leqslant s < T. \tag{55}$$

Aguirregabiria and Mira (2010) define a more general set of assumptions for Rust's model. We give here the equivalent assumptions on (21) within our framework:

- Additive Separability (AS): no serial correlation in individual utilities ($\alpha_{in}^u = 0$ and $\lambda_\varepsilon = 1$);
- i.i.d. unobservables (IID): random portion of error term ($\varepsilon_{ins}$) is distributed i.i.d. (as exemplified in (23)).
- Conditional independence of future $x$ (CI-X): no serial correlation in individual anticipation, and no variance increase with longer-term prediction ($\alpha_n^x = 0$ and $\lambda_\nu = 1$ in (19));
- Conditional independence of $y$ (CI-Y): in the formulation in this chapter, the *payoff variables* are included in $x$, therefore this assumption is satisfied by the assumptions for CI-X;
- CLOGIT: random portion of error term ($\varepsilon_{ins}$) has a Type 1 GEV distribution (as tested in (23)); and
- Discrete support of $x$ (DIS): observed explanatory variables $x_{nt}$ are finite.

As we highlight in Section 4.3, there is typically a high computational burden associated with estimating dynamic discrete choice models. Nevertheless, there are many examples of successful applications in the literature. Some studies use models similar to that of Rust (1987), while others relax certain of the aforementioned assumptions (e.g., Eckstein & Wolpin, 1989; Erdem & Keane, 1996; Keane & Wolpin, 1997).

The focus of the application in Rust (1987) could be viewed as closer to that of inverse optimization than analyzing and predicting choice behaviour. Indeed, the focus lies on a

[7] The utility of not making the replacement can be fixed to zero as only differences in utility matter.

single individual and the optimization problem (bus engine replacement) he solves as part of his work. Works aimed at analyzing and predicting the choice behaviour of a population deal with applications in various domains. In the following we briefly describe a few examples.

Karlstrom et al. (2004) study how the pension system affects the retirement choice of blue-collar workers in Sweden. Each year between the age of 50 and 70, they model forward-looking individuals' choice of retiring or not. It hence corresponds to an optimal stopping problem, similar to the one of Rust (1987).

Dynamic discrete choice models are well suited to model the choice behaviour of durable goods. A prominent example is car ownership choice. Gillingham et al. (2015) propose a model of households' car buy and sell decisions as well as the car owners' usage. Equilibrium prices in the used-car market are endogenous to the model which is estimated based on Danish register data covering all Danish households and cars over more than a decade.

Another application of dynamic discrete choice models for durable goods is to model consumer stockpiling. Ching and Osborne (2020) investigate the household purchase behaviour of laundry detergent across multiple product brands and sizes. The model includes distributed parameters to account for unobserved heterogeneity in discount factors and price coefficients.

There are many parallels between structural economics (SE) dynamic discrete choice models and inverse reinforcement learning (IRL) algorithms (Ng & Russell, 2000). IRL aims at extracting a reward function from a set of observed optimal trajectories, and is hence similar to the forward-looking dynamic discrete choice model presented in this chapter. Despite these parallels, the literature on IRL has to a large extent evolved separately from that of SE. Iskhakov et al. (2020) discuss contrasts and synergies between the two fields. The authors note that the methods used in each field are quite different. Notably, IRL does not pose the problem as one of parameter estimation. This is partly due to the difference between the intended applications: IRL is focused on prediction while SE is concerned with inference and *counter factual* prediction.

## 6.2 Markov models

The Markov model is typically applied in the literature to describe myopic behaviour, where the decision maker evaluates only the utility at time $t$ without taking into account any future consequences from their choice. This can be achieved by fixing the discount parameter $\rho_n$ in (22) to zero. The global utility $U'_i$ in (13) is then equal to the instantaneous utility $U_i$ (and, by extension, $V'_i = V_i$). The utility function for the Markov model (33) thus becomes:

$$U'_i(\tilde{x}_{nt}) = V_i(\tilde{x}_{nt}) + \eta_i y_{n,t-1} + \alpha^U_{in} + \varepsilon_{int}, \ t \leqslant T, \tag{56}$$

where $V_i$ is a deterministic function of the observable explanatory variables $x_{nt}$.

Wooldridge (2005) uses a probit Markov model to investigate the persistence of working union membership, where the decision variable $y_{nt}$ is a binary variable that represents the decision of individual $n$ to be a member of a union in year $t$ or not, and the only time dependent explanatory variable $x_{nt}$ is a binary variable that represents if the individual $n$ is married at time $t$ or not. This model can be obtained by applying the following further assumptions on (33):

1. the error terms are normally distributed (to give the probit model) with constant variance $\sigma_\varepsilon^2$: $\varepsilon_{int} \sim N(0, \sigma_\varepsilon^2)$; and
2. the agent effects are given by $\alpha_{in}^U = a_U y_{n,t_b} + b_U^T x_n + \zeta_{in}^U$, where $\zeta_{in}^U \sim N(0, \sigma_\alpha^2)$.

For the specific example in the paper, the instantaneous utility is given by $V(\tilde{x}_{nt}) = \beta_x x'_{nt} + c_t + \beta_0$, where $\beta_0$ is a single constant and $c_t$ is a constant for each time period in the dataset, to be estimated from the data. Similar to Rust's model of bus engine replacement, the $i$ subscript can be dropped as the decision variable is binary. This gives the final form for the global utility of

$$U'(\tilde{x}_{nt}) = \beta_x x'_{nt} + c_t + \beta_0 + \eta y_{n,t-1} + \alpha_U y_{nt_b} + b_U^T x_n + \zeta_n^U + \varepsilon_{nt}. \quad (57)$$

Note that the constant term in the agent effects formula in Wooldridge's model is included in $\beta_0$ in this formulation.

There have been several other applications of Markov models to investigate dynamic choice situations which make use of Wooldridge's correction method for the initial condition problem. Muûls and Pisu (2009) estimate models of organisational level import and export decisions for all Belgian companies over an eight-year period. Separate models are estimated for export and import. In each case the decision variable represents the binary decision to export (or import) or not in a given year.

As with forward-looking models, Markov models have also been applied to investigate car ownership behaviour. Nolan (2010) estimate a dynamic probit model of household car ownership in Ireland using six-years of longitudinal household survey data. As with the application of Muûls and Pisu, a binary decision variable is used (whether a household owns a car during the survey period or not).

Wooldridge's correction method has also been applied to multiclass problems. For example, Danalet et al. (2016) estimate a Markov model for a catering location choice problem on a university campus with 21 alternatives. The model makes use of WiFi traces to calculate additional explanatory variables, such as the distance from previous activity locations. Furthermore, the model makes use of multiple separate lagged choices from the previous period, namely the location choice for morning and lunch periods in the previous day.

There have been applications of Markov models which relax the Markov assumption, and allow for higher order lagged variables in the utility specification. For example, Bogers et al. (2007) model the effect of learning in route choice, and include a weighted average of the previous 10 choices in the utility specification.

Whilst not covered explicitly in this chapter, Markov models have also been applied to estimate ordinal models. For example, Contoyannis et al. (2004) estimate an ordered probit model of self-assessed health status using data from the a household panel survey from the UK.

## 6.3 Hidden Markov models

Applications of the hidden Markov model for dynamic choice can be grouped into two categories. The first category are models which include the change in a continuous *autoregessive latent variable* in the utility specification. The second category of models map the transitions between a finite number of discrete *latent classes*, each with a different set of model parameters. We provide first the assumptions needed to derive an example of the former, and then discuss further examples of both approaches.

Heiss (2008) models the self-reported health status of survey respondents in the USA. An ordered logit model is used to predict the response within a five-point scale from poor to excellent. The latent continuous agent effects are allowed to vary over time, dependent on their previous value. The resulting model is hence a hidden Markov model. It can be derived from (56) through the following assumptions/modifications:

1. the agent effects/serial correlation $\alpha^{u}_{in}$ are allowed to vary over time according to the pdf $f_{\alpha^u_i}(\alpha^{u}_{int}|\tilde{x}_{nt}, \alpha^{u}_{in,t-1})$, and

2. the previous choice does not affect the utility, so that $\eta_i = 0 \; \forall \; i$.

For the specific example in the paper, the pdf $f_{\alpha^u_i}(\alpha^{u}_{int}|\tilde{x}_{nt}, \alpha^{u}_{in,t-1})$ is a normal stationary auto-regressive process of order one, independent of $\tilde{x}_{nt}$

$$\alpha^{u}_{int} = \kappa\alpha^{u}_{in,t-1} + \varepsilon^{\alpha}_{int} \tag{58}$$

where $\kappa$ is a correlation parameter to be estimated from the data and $\varepsilon^{\alpha}_{int}$ is normally distributed. Furthermore, the ordered logit model is for only one aspect (health status) and so the $i$ subscript can be dropped. This gives the following form of the global utility:

$$U'(\tilde{x}_{nt}) = V(\tilde{x}_{nt}) + \kappa\alpha^{u}_{n,t-1} + \varepsilon'_{nt}, \; t \leqslant T. \tag{59}$$

Heiss et al. (2010) build on this work to investigate subscription to basic health insurance (Medicare) in the USA using annual health survey data for respondents aged sixty-five and over. The decision variable is a binary choice of whether to enroll in the Medicare program in the survey year (or not). Enrollment is assumed to be a permanent decision, such that once a person has a plan (i.e. if $y_{nt} = 1$) they will then keep the plan for all future time periods. A latent continuous variable which measures the *health capital* is included in the utility specification, based on an autoregressive latent robustness. The value of the health capital is estimated based on its structural relations with the survival indicator, self-reported health status, and pharmacy bills.

As well as latent continuous variables, the hidden Markov model can be used to model changes between a finite number of discrete latent classes, each with their own utility specifications or parameter values. Netzer et al. (2008) model the binary choice of alumni donating (or not) in a survey year based on the respondent's latent relationship state with their alma mater, which is allowed to change over time. Models with different numbers of states between two and four are tested. This approach has also been used to investigate multiclass problems. For example, Xiong et al. (2015) investigate an individual mode-choice problem out of five possible travel modes using panel data over a 10-year period, based on switching between two latent preference states.

There has also been work to relax the Markov assumption in dynamic choice models by allowing for higher order lagged variables in the utility specification. Xiong et al. (2018) investigate the use of second order $(t-2)$ and third order $(t-3)$ lagged variables in a model of dynamic car ownership. Second and third-order hidden Markov with two latent classes are compared against first-order models with two/three latent classes. The second-order model with two latent states was found to have the lowest Bayesian Information Criteria (BIC).

# 7 Conclusion

Dynamic choice models in the literature typically belong to one of two categories:

1. forward-looking models based on dynamic programming formulations using Bellman's principle of optimality, or

2. models to describe habitual behaviour and learning models based on the Markov assumption that assume myopic behaviour.

In this chapter, we analyse the dynamic choice problem, both from the point of view of the modeller and of the analyst, to derive a general parametric dynamic choice model based on first principles. This general model extends the state of practice by (i) unifying forward-looking models and habitual behaviour and learning models under a single general framework; (ii) specifically discussing the Markov assumption in the anticipation of future explanatory variables and habitual behaviour and learning; (iii) including agent effects in both the utility function and the anticipation of future explanatory variables; and (iv) accounting for variance inflation in the error terms in the future utility and anticipation of future explanatory variables as the prediction interval (i.e. $s - t$) increases.

We use the general model to show how different types of dynamic choice models in the literature can be derived through simple assumptions on the model parameters. We derive a specific example for each type of model, and then introduce several further examples of applications of each model type in the literature. This approach clearly illustrates the differences between dynamic models used in the literature through their implied assumptions.


# Acknowledgement

We are grateful to Moshe Ben-Akiva for valuable discussions that helped us improving the quality of this chapter. We also express our gratitude to Daniel McFadden, whose lecture notes have been an important source of inspiration.


# References


Aguirregabiria, V., & Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, *156*(1), 38–67. https://doi.org/10.1016/j.jeconom.2009.09.007

Bellman, R. (1952). On the Theory of Dynamic Programming. *Proceedings of the National Academy of Sciences*, *38*(8), 716–719. https://doi.org/10.1073/pnas.38.8.716

Bertsekas, D. P. (2017). *Dynamic programming and optimal control* (4th ed., Vol. I & II). Athena scientific Belmont, MA.

Bogers, E. A. I., Bierlaire, M., & Hoogendoorn, S. P. (2007). Modeling Learning in Route Choice. *Transportation Research Record*, *2014*(1), 1–8. https://doi.org/10.3141/2014-01

Ching, A. T., & Osborne, M. (2020). Identification and Estimation of Forward-Looking Behavior: The Case of Consumer Stockpiling. *Marketing Science*, *39*(4), 707–726. https://doi.org/10.1287/mksc.2019.1193

Cirillo, C., Xu, R., & Bastin, F. (2015). A Dynamic Formulation for Car Ownership Modeling. *Transportation Science*, *50*(1), 322–335. https://doi.org/10.1287/trsc.2015.0597

Contoyannis, P., Jones, A. M., & Rice, N. (2004). The dynamics of health in the British Household Panel Survey. *Journal of Applied Econometrics*, *19*(4), 473–503. https://doi.org/10.1002/jae.755

Danalet, A., Tinguely, L., de Lapparent, M., & Bierlaire, M. (2016). Location choice with longitudinal WiFi data. *Journal of Choice Modelling*, *18*, 1–17. https://doi.org/10.1016/j.jocm.2016.04.003

Eckstein, Z., & Wolpin, K. I. (1989). The specification and estimation of dynamic stochastic discrete choice models: A survey. *The Journal of Human Resources*, *24*(4), 562–598. https://doi.org/10.2307/145996

Erdem, T., & Keane, M. P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing science*, *15*(1), 1–20. https://doi.org/10.1287/mksc.15.1.1

Fosgerau, M., Frejinger, E., & Karlstrom, A. (2013). A link based network route choice model with unrestricted choice set. *Transportation Research Part B: Methodological*, *56*, 70–80. https://doi.org/10.1016/j.trb.2013.07.012

Gillingham, K., Iskhakov, F., Munk-Nielsen, A., Rust, J., & Schjerning, B. (2015). *A dynamic model of vehicle ownership, type choice, and usage* (Working paper).

Greene, W. H. (2001, January 1). *Fixed and Random Effects in Nonlinear Models* (Economics Working Papers EC-01-01). New York University. NY, USA.

Guevara, C. A., Tang, Y., & Gao, S. (2018). The initial condition problem with complete history dependency in learning models for travel choices. *Transportation Research Part B: Methodological*, *117*, 850–861. https://doi.org/10.1016/j.trb.2017.09.006

Heiss, F. (2008). Sequential numerical integration in nonlinear state space models for microeconometric panel data. *Journal of Applied Econometrics*, *23*(3), 373–389. https://doi.org/10.1002/jae.993

Heiss, F., McFadden, D., & Winter, J. (2010). Mind the gap! Consumer perceptions and choices of Medicare Part D prescription drug plans, In *Research findings in the economics of aging*. University of Chicago Press.

Hotz, V. J., & Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, *60*(3), 497–529. https://doi.org/10.2307/2298122

Hotz, V. J., Miller, R. A., Sanders, S., & Smith, J. (1994). A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies*, *61*(2), 265–289. https://doi.org/10.2307/2297981

Imai, S., Jain, N., & Ching, A. (2009). Bayesian Estimation of Dynamic Discrete Choice Models. *Econometrica*, *77*(6), 1865–1899. https://doi.org/10.3982/ecta5658

Iskhakov, F., Rust, J., & Schjerning, B. (2020). Machine learning and structural econometrics: Contrasts and synergies. *The Econometrics Journal*, *23*(3), 81–124. https://doi.org/10.1093/ectj/utaa019

Julier, S. J., & Uhlmann, J. K. (1997, July 28). New extension of the Kalman filter to nonlinear systems, In *Signal Processing, Sensor Fusion, and Target Recognition VI*. AeroSense '97, Orlando, FL, USA, International Society for Optics and Photonics. https://doi.org/10.1117/12.280797

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, *82*(1), 35–45. https://doi.org/10.1115/1.3662552

Karlstrom, A., Palme, M., & Svensson, I. (2004). A dynamic programming approach to model the retirement behaviour of blue-collar workers in Sweden. *Journal of Applied Econometrics*, *19*(6), 795–807. https://doi.org/10.1002/jae.798

Keane, M. P., & Wolpin, K. I. (1994). The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence. *The Review of Economics and Statistics*, *76*(4), 648–672. https://doi.org/10.2307/2109768

Keane, M. P., & Wolpin, K. I. (1997). The Career Decisions of Young Men. *Journal of Political Economy*, *105*(3), 473–522. https://doi.org/10.1086/262080

Manski, C. F. (1973). *The analysis if qualitative choice.* (Thesis). Massachusetts Institute of Technology. Boston, MA, USA. Retrieved February 24, 2021, from https://dspace.mit.edu/handle/1721.1/13927

Manski, C. F. (1977). The structure of random utility models. *Theory and decision*, *8*(3), 229. https://doi.org/10.1007/bf00133443

Muûls, M., & Pisu, M. (2009). Imports and Exports at the Level of the Firm: Evidence from Belgium. *The World Economy*, *32*(5), 692–734. https://doi.org/10.1111/j.1467-9701.2009.01172.x

Netzer, O., Lattin, J. M., & Srinivasan, V. (2008). A Hidden Markov Model of Customer Relationship Dynamics. *Marketing Science*, *27*(2), 185–204. https://doi.org/10.1287/mksc.1070.0294

Ng, A. Y., & Russell, S. (2000). Algorithms for Inverse Reinforcement Learning, In *Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000)*, Stanford University, Stanford, CA, USA, Morgan Kaufmann. https://doi.org/10.5555/645529.657801

Nolan, A. (2010). A dynamic analysis of household car ownership. *Transportation Research Part A: Policy and Practice*, *44*(6), 446–455. https://doi.org/10.1016/j.tra.2010.03.018

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, *55*(5), 999. https://doi.org/10.2307/1911259

Rust, J., & Phelan, C. (1997). How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets. *Econometrica*, *65*(4), 781–831. https://doi.org/10.2307/2171940

Su, C.-L., & Judd, K. L. (2012). Constrained Optimization Approaches to Estimation of Structural Models. *Econometrica*, *80*(5), 2213–2230. https://doi.org/10.3982/ecta7925

Västberg, O. B., Karlström, A., Jonsson, D., & Sundberg, M. (2019). A Dynamic Discrete Choice Activity-Based Travel Demand Model. *Transportation Science*, *54*(1), 21–41. https://doi.org/10.1287/trsc.2019.0898

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, *20*(1), 39–54. https://doi.org/10.1002/jae.770

Xiong, C., Chen, X., He, X., Guo, W., & Zhang, L. (2015). The analysis of dynamic travel mode choice: A heterogeneous hidden Markov approach. *Transportation*, *42*(6), 985–1002. https://doi.org/10.1007/s11116-015-9658-2

Xiong, C., Yang, D., & Zhang, L. (2018). A high-order hidden markov model and its applications for dynamic car ownership analysis. *Transportation Science*, *52*(6), 1365–1375. https://doi.org/10.1287/trsc.2017.0792

Zimmermann, M., & Frejinger, E. (2020). A tutorial on recursive models for analyzing and predicting path choice behavior. *EURO Journal on Transportation and Logistics*, *9*(2), 100004. https://doi.org/10.1016/j.ejtl.2020.100004