Identification of Time Preferences in Dynamic Discrete Choice Models: Exploiting Choice Constraints

Ulrich C. Schneider*†

Abstract

This paper studies sufficient conditions for the identification of the exponential discount factor in dynamic discrete choice models, with a focus on models in which decision makers might only be able to choose from a subset of all potential alternatives. The presented exclusion condition requires variation in the probability of being "choice constrained". Variation in such constraining probabilities that does not affect utilities or transition probabilities leads to point identification of the exponential discount factor. The derived identifying equation aligns with economic intuition. Some examples are provided.

JEL: C14, C23, C61

Keywords: discount factor; identification; dynamic discrete choice

*Freie Universität Berlin, u.schneider@fu-berlin.de, Tel. +49 30 838 69494, Boltzmannstr. 20, 14195 Berlin, Germany.

† *Acknowledgments*: This paper is an adapted version of Chapter 1 in the PhD thesis of the corresponding author (see Schneider 2019). Funding by the German Federal Ministry of Labour and Social Affairs (BMAS), within the framework "Interdisciplinary Social Policy Research" (FIS) is gratefully acknowledged. The author thanks Fedor Iskhakov and Arturas Juodis for guidance and support.

1 Introduction

Dynamic discrete choice models are used to estimate the behaviour of economic agents and analyse counterfactual policies in numerous fields. The discount factor is crucial in these models, as it determines the reactions to expected future events. In empirical applications, researchers are often compelled to fix the discount factor because it is challenging to identify the parameter from the observed choice probabilities without further restrictions or without imposing a sufficient structure (see Magnac and Thesmar 2002; Rust 1994). Time preferences appear to be stable over time (Meier and Sprenger 2015) but vary substantially across contexts and populations (Frederick, Loewenstein, and O'Donoghue 2002). Thus, setting the discount factor to a predetermined value can lead to incorrect inference and misleading counterfactual analysis.

This paper provides identification of the exponential discount factor by exploiting the fact that decision makers are often constrained in their choices. Exogenous variation in the probability of being constrained in one's choice point identifies the discount factor in dynamic discrete choice models. The intuition behind the identification result relies on the known insight that variation in future expected values that does not affect instantaneous utilities can be exploited to identify time preferences.

Economic examples of such constraints are plentiful. Potential car buyers can only choose among the set of alternatives available at the dealership. A maintenance supervisor wishing to replace a bus engine might be unable to do so because no engines are currently in stock. An unemployed individual might not be able to start working if he or she fails to receive a job offer. A firm might not be able to merge with another company or enter a protected market due to regulatory disapproval.

In many of these contexts, the constraint on decision makers comes either from a regulatory authority or the opposite side of the market. Thus, probabilities that determine if agents are constrained in their decision are ideally suited to provide variation that does not directly impact the decision makers' instantaneous utilities but only their future choices. For example, a labor demand shock influences the probability that a job searching individual receives a job offer. It should not directly affect the individual's human capital stock, age, or other state variables. If researchers observe individuals in two labor markets that differ only with respect to the job offer probability, the exponential discount factor can be identified.

Other models might be adapted to fit the presented framework. For instance, consider the classic bus engine replacement example of Rust (1987). Assume that one bus engine is in stock. Once a bus engine is replaced, Harold Zurcher has to order a new engine. He is constrained in his choice if a new bus engine cannot be delivered immediately, for example, due to a strike or other supply shocks. As a result, Harold Zurcher no longer has the option to replace a bus engine for two consecutive periods. This variation would be sufficient to point identify the exponential discount factor since it neither directly affects the transition probabilities of the buses' mileage nor Harold Zurcher's instantaneous utilities.

The framework presented here also permits the probability of being constrained to be zero or one. Thus, settings in which choice sets vary deterministically are nested. For the identification of β, it is only crucial that there is some variation in the probability of being constrained. Such variation can also be a change from zero to one or vice versa.

The identification of time preferences does not require the normalization of one alternative's utility. Although such a normalization is a standard approach in empirical applications, it can cause misleading counterfactual policy simulations in some cases (see, e.g., Norets and Tang 2014; Kalouptsidi, Scott, and Souza-Rodrigues 2019). Furthermore, the identification of the discount factor does not exploit the stationarity of the model. Contrary to the identification of the discount factor, the identification of the instantaneous utilities requires both the normalization of one alternative's utility and stationarity.

The literature discussing dynamic discrete choice models (e.g., Abbring 2010; Arcidiacono and Miller 2020; Srisuma 2015; Hu and Sasaki 2018; Chen 2017) has examined various aspects of these models, with few papers providing conditions to identify time preferences. Nevertheless, the identification of time preferences is pivotal for the identification of alternative specific utilities. Magnac and Thesmar (2002) demonstrate that it is not possible to nonparametrically identify utilities if the discount factor is not known.[1]

1. Aguirregabiria and Suzuki (2014) and Komarova et al. (2018) show that by imposing a certain structure,

Conversely, the two authors show that once the discount factor is identified and given the distribution of preference shocks and the normalization of one utility, all other utilities can be uniquely determined. This identification result can be adjusted to apply to the setting with choice constraints and is discussed in Section 4.

In their seminal paper, Magnac and Thesmar (2002) also examine how the exponential discount factor can be recovered from the observed choices. Their derived exclusion restriction requires two states that have different expected streams of utilities for some choices but equal expected streams for at least one choice. These states should also provide the same instantaneous utilities. The requirements on future expected utility streams increase the difficulty of finding such states in empirical applications.

The exclusion restriction presented in this paper is formulated on instantaneous utilities and does not directly include future value functions. This simplifies the interpretation and facilitates identifying situations that satisfy the exclusion restriction in empirical applications.

Abbring and Daljord (2020) extend the identification approach of Magnac and Thesmar (2002). They also develop an exclusion restriction on instantaneous utilities. Their exclusion restriction requires two states that have equal instantaneous utilities but different transition probabilities. In an infinite horizon model, the identifying equation includes an infinite geometric sum. As a result, time preferences are only set identified. When excluding discount factors that are close to one, the set consists of a finite number of discount factors. Abbring and Daljord (2020) discuss several examples that allow for a reduction in the number of potential solutions, for example, by relying on the concept of finite dependence (see Arcidiacono and Miller 2019) or using multiple exclusion restrictions.[2]

Another approach to identify the discount factor is provided by Komarova et al. (2018), who assume a certain structure of the model to aid identification. They provide identification results for linear-in-parameter models. Similar to the result of this paper, they show that time preferences can be identified without normalizing the utility of one alternative. The approach presented here imposes more structure on the choice sets from which decision makers can

it is possible to identify payoff or cost functions without knowing the discount factor.

2. The work of Levy and Schiraldi (2021) is related to this paper. While they focus on the history of made choices, I consider variation in the probability of being constrained.

choose. No structure is imposed on the utilities or payoff functions.

The remainder of the paper is structured as follows: Section 2 introduces the model, and Section 3 discusses the identification of the exponential discount factor. Section 4 provides sufficient conditions to identify instantaneous utilities for a framework allowing decision makers to be constrained in their choices. In some contexts, the probabilities that limit agents' choices cannot be observed. Section 5 discusses sufficient conditions that allow one to identify these probabilities. Similarly, in situations in which the researcher is unable to observe from which choice set the agents chose, the observed choices are a mix of preferences and the likelihood of being choice constrained. For such cases, Section 6 discusses how choice probabilities solely based on preferences can be recovered. Finally, Section 7 concludes the paper.

2 Model

In each period t, where time is discrete with an infinite horizon, a decision maker has to choose an alternative from a finite set of alternatives. In a given period, the agent faces either the *general* choice set, denoted by $\mathcal{D}_G$, or a *constrained* choice set, denoted by $\mathcal{D}_R(d)$.

Assumption 1. (Choice Sets). *The general choice set is* $\mathcal{D}_G = \{1, \ldots, K\}$*, with* $K > 1$*. For each alternative* $d \in \mathcal{D}_G$*, the respective constrained choice set is defined as* $\mathcal{D}_R(d) \subset \mathcal{D}_G$ *with* $\mathcal{D}_R(d) \neq \varnothing$*. The researcher knows which alternatives are included in each constrained choice set.*

Note that there are K constrained choice sets, which could all be equal. The choice set the agents must choose from depends on the last period's choice and the state $x \in \mathcal{X}$, where $\mathcal{X}$ contains a finite number of J elements (finite support assumption).

Assumption 2. (Constraining probabilities[3]). *With probability* $\pi(d_t, x_t)$*, agents must choose from* $\mathcal{D}_R(d_t)$ *in period* $t+1$*. With probability* $1 - \pi(d_t, x_t)$*, agents must choose from* $\mathcal{D}_G$*.*

3. Note that the constraints discussed here are not identical to those in McFadden (1978). McFadden (1978) introduces constraints to reduce the computational burden in problems with a large number of alternatives. Constrained choice sets should also not be interpreted as *consideration sets*, as for instance, in Shocker et al. (1991) or Goeree (2008). Consideration sets exclude alternatives that agents are not aware of and, thus, do not consider. This paper assumes that agents are fully aware of potential constraints and base their decisions at least partly on future probabilities of being constrained.

Furthermore, $\exists(d_t, x_t) \in \mathcal{D}_G \times \mathcal{X}$ *such that* $\pi(d_t, x_t) > 0$.

Choosing alternative d_t, given the state x_t, provides agents with the instantaneous utility $u^\star\left(d_t, x_t, \eta_{t,d_t}\right)$, where η_{t,d_t} denotes an alternative specific preference shock.

Assumption 3. (Distribution of preference shocks). *The preference shock* $\eta_{t,d_t} \overset{i.i.d.}{\sim}$ *mean-zero, type-I extreme value.* η_{t,d_t} *cannot be observed by the researcher.*

The identification results extend to any other known continuous distribution on $\mathbb{R}^K$. There are two reasons for explicitly choosing the type-I extreme value distribution. First, it eases the exposition. Second, it can be found in many empirical applications.[4] The vector $\boldsymbol{\eta}_t = \{\eta_{t,1}, \ldots, \eta_{t,K}\}$ contains the preference shocks for all alternatives in a given period.

Another common assumption is that the instantaneous utility is additively separable in the preference shock.

Assumption 4. (Additive separability). *Utility* $u^\star\left(d_t, x_t, \eta_{t,d_t}\right) = u(d_t, x_t) + \eta_{t,d_t}\ \forall(d_t, x_t, \eta_{t,d_t}) \in \mathcal{D}_G \times \mathcal{X} \times \boldsymbol{\eta}_t$.

The next assumption is on the development of the state space. The probability of observing state x_{t+1} in the next period depends on both the choice d_t and the state x_t of the current period but is independent of any preference shock

Assumption 5. (Conditional independence assumption). *State* x_{t+1} *occurs with the transition probability* $q(x_{t+1} | d_t, x_t)$. *It is assumed that neither the transition nor the constraining probabilities depend on any preference shocks* $\boldsymbol{\eta}_t\ \forall t$.

Finally, an assumption about the stationarity of the model is made. Note that the identification of the discount factor does not exploit the stationarity of the model,[5] but identification of the instantaneous utilities does.

Assumption 6. (Stationarity). *The model is stationary.*

4. Recent studies on the estimation of models including a combination of discrete and continuous choices even add a type-I extreme value shock to better handle potential kinks in the value functions (see Iskhakov et al. 2017).

5. Thus, the identification discussion can also be adjusted to suit a finite horizon model.

The order of events in each period t is the following. First, the choice set is determined, and agents observe state x_t and preference shocks $\boldsymbol{\eta}_t$. Second, agents choose one alternative d_t. Finally, the agents collect the instantaneous utility $u^\star\left(d_t, x_t, \eta_{t,d_t}\right)$, and the period ends.

In period t, agents choose the alternative that maximizes their total expected discounted stream of instantaneous utilities $v^\star\left(d_t, x_t, \eta_{t,d_t}\right)$. Assumptions 4 and 5 imply $v^\star\left(d_t, x_t, \eta_{t,d_t}\right) = v\left(d_t, x_t\right) + \eta_{t,d_t}$, with

$$\begin{aligned} v\left(d_t, x_t\right) = u(d_t, x_t) + \beta \sum_{\substack{x_{t+1}\\ \in \mathcal{X}}} \Bigg[& \left(1 - \pi\left(d_t, x_t\right)\right) \mathbb{E}\left[\max_{d_{t+1}\in \mathcal{D}_G} \left\{v\left(d_{t+1}, x_{t+1}\right) + \eta_{t+1,d_{t+1}}\right\}\right] \\ & + \pi\left(d_t, x_t\right) \mathbb{E}\left[\max_{d_{t+1}\in \mathcal{D}_R(d_t)} \left\{v\left(d_{t+1}, x_{t+1}\right) + \eta_{t+1,d_{t+1}}\right\}\right]\Bigg] q\left(x_{t+1} \mid d_t, x_t\right), \end{aligned} \tag{1}$$

where β denotes the discount factor. Under assumption 3, equation (1) can be expressed as

$$\begin{aligned} v\left(d_t, x_t\right) = u(d_t, x_t) + \beta \sum_{\substack{x_{t+1}\\ \in \mathcal{X}}} \Bigg[& \left(1 - \pi\left(d_t, x_t\right)\right) \ln\left(\sum_{d_{t+1}\in \mathcal{D}_G} \exp\left(v\left(d_{t+1}, x_{t+1}\right)\right)\right) \\ & + \pi\left(d_t, x_t\right) \ln\left(\sum_{d_{t+1}\in \mathcal{D}_R(d_t)} \exp\left(v\left(d_{t+1}, x_{t+1}\right)\right)\right)\Bigg] q\left(x_{t+1} \mid d_t, x_t\right). \end{aligned} \tag{2}$$

This paper concentrates on the identification of the time preference parameter β in equation (1) because it is essential to the identification of the model. Suppose that data are available such that all transition, constraining, and genuine choice probabilities are known. Then, the model is point identified if and only if utilities can be uniquely determined from these probabilities. As discussed in Section 4, an adapted version of Proposition 2 of Magnac and Thesmar (2002) is fulfilled: In a stationary infinite horizon model and given the distribution of preference shocks and the normalization of one alternative's utility, all other utilities depend on the discount factor β. Thus, without knowing β, utilities cannot be uniquely determined, and the model is not identified.[6]

The discussion concerning identification distinguishes between *composite* and *genuine* choice probabilities. If the data are not informative about the choice set agents face when

6. Komarova et al. (2018) show that under a linear-in-parameter specification, it is possible to determine utilities without knowing the discount factor.

making their choices, the researcher can only observe the composite choice probabilities. These reflect a combination of preferences and possible choice constraints. The composite choice probabilities are denoted by $\mathrm{KPr}\left(d_t \,|d_{t-1}, x_{t-1}, x_t,\right)$. Genuine choice probabilities describe the probability that agents choose alternative d_t from a specific choice set $\mathcal{D} \in \{\mathcal{D}_G, \mathcal{D}_R\left(d_{t-1}\right); d_{t-1} \in \mathcal{D}_G\}$, given state x_t. They exclusively reflect preferences over the alternatives included in the respective choice set. Genuine choice probabilities are denoted by $\Pr\left(d_t \,|\mathcal{D}, x_t\right)$. Section 6 discusses the identification of genuine choice probabilities if they are not observed directly.

For ease of exposition, I drop the subscript t henceforth and denote the variables of period $t+1$ by a prime. The following sections also assume that the constraining and genuine choice probabilities are known. Section 5 discusses the identification of constraining probabilities when they are not known.

3 Identification of the exponential discount factor

I assume that $\{\Pr\left(d\,|\mathcal{D}_G, x\right),\ q\left(\cdot\,|d, x\right),\ \pi\left(d, x\right);\ (d, x) \in \mathcal{D}_G \times \mathcal{X}\}$ are known for at least two consecutive periods for the same group of decision makers.[7]

As noted in previous literature (see for example Fang and Wang 2015; Abbring and Daljord 2020), a fruitful strategy to identify time preferences is to find variation that affects the future but not the current period. In the presented model, such variation is given when the constraining probabilities vary exogenously.[8]

Assumption 7. (Exclusion Restriction). *There exist two different states $x^A, x^B \in \mathcal{X}$ and two different alternatives $\ell, r \in \mathcal{D}_G$ such that*

(1) $u\left(\ell, x^A\right) = u\left(\ell, x^B\right)$ *and* $u\left(r, x^A\right) = u\left(r, x^B\right)$*;*

(2) $q\left(x' \,\middle|d, x^A\right) = q\left(x' \,\middle|d, x^B\right)$ *for* $d \in \{\ell, r\}$*, and* $x' \in \mathcal{X}$*; and*

7. Note that for a given $(d, x) \in \mathcal{D}_G \times \mathcal{X}$, $\Pr\left(d\,|\mathcal{D}_G, x\right)$ identifies $\Pr\left(d\,|\mathcal{D}, x\right)$, $\forall \mathcal{D} \in \{\mathcal{D}_R\left(d\right); d \in \mathcal{D}_G\}$. For details, see Lemma 1 of Section 6.

8. Note that if transition probabilities are also affected, this might provide an additional source for identification (see Abbring and Daljord 2020).

(3) $\pi\left(\ell, x^A\right) \neq \pi\left(\ell, x^B\right)$.

The exclusion restriction is formulated for two states and two alternatives that must fulfil three conditions. First, each alternative must provide the same instantaneous utility for both states. Second, for each alternative, the transition probabilities must be equal for both states. Third, for at least one alternative, the constraining probabilities must differ between the two states.

In other words, a state is required that does not affect the instantaneous utility or the transition probabilities but does affect the constraining probabilities. In a labor supply model, the exclusion restriction is fulfilled when there is a labor demand shock that only affects job offer probabilities. Other state variables, e.g., human capital, should not be affected by a demand shock.[9]

Another example in the context of labor supply is provided by Lalive (2008). He describes the situation in Austria where the length of unemployment benefits varies across regions. The following model would satisfy the exclusion restriction: the general choice set consists of the three alternatives *employment*, *unemployment*, and *out of the labor force*. While individuals in region A are entitled to two periods of unemployment, individuals in region B are entitled to three. Individuals who already spent two periods in unemployment have different restriction probabilities across regions. It seems reasonable that the region itself does not affect the transition probabilities, or the utility itself.

Conlone and Mortimer (2013) provide a framework for product demand. They exploit the availability of products in vending machines to estimate the demand for products. Variation in the refilling schedules for different machines would result in different constraining probabilities without influencing other state variables of customers.

Harford (2005) identifies (unexpected) regulatory changes as one factor that drives mergers and acquisitions. A careful selection of firms that temporarily operate under different regulations can provide variation that fulfils assumption 7.

9. Haan, Haywood, and Schneider (2017) use such variation for women on maternity leave. Women who are not working while on maternity leave have the right to return to their previous job for a limited time, which affects their job offer probability. Reforms that change the duration of maternity generate exogenous variation in the job offer probabilities for different women on maternity leave.

In general, constraining probabilities can be interpreted as external factors that only affect agents' choice sets. If the supply side is modelled, constraining probabilities can be seen as frictions coming from the demand side and vice versa. Furthermore, regulations and reforms that influence the potential choice set can be exploited for identification.

For the identification of the exponential discount factor, a rank condition is required. The rank condition can be more compactly stated if we define for a given $(d, x) \in \mathcal{D}_G \times \mathcal{X}$ the vector $\boldsymbol{q}(d, x)$, which has size K. The k^{th} element of $\boldsymbol{q}(d, x)$ is $q(x_k | d, x)$ with $x_k \in \mathcal{X}$. Given two states x^A and x^B that satisfy the exclusion restriction stated in 7, the rank condition is given by

Assumption 8. (Rank condition).

$$\left(\pi\left(\ell, x^B\right) - \pi\left(\ell, x^A\right)\right) \boldsymbol{q}\left(\ell, x^A\right)^{\top} \boldsymbol{\Theta}(\ell) - \left(\pi\left(r, x^B\right) - \pi\left(r, x^A\right)\right) \boldsymbol{q}\left(r, x^A\right)^{\top} \boldsymbol{\Theta}(r) \neq 0, \tag{3}$$

where $\boldsymbol{\Theta}(d)$ is a vector of size J and the superscript $\top$ denotes the transpose. For a freely chosen alternative $d' \in \mathcal{D}_R(d)$, the jth element of $\boldsymbol{\Theta}(d)$ is

$$\begin{aligned}\boldsymbol{\Theta}(d)_{[j]} = &\left(\mathbb{E}\left[\max_{i' \in \mathcal{D}_G} \left\{v\left(i', x'\right) + \eta'_{i'}\right\}\right] - v\left(d', x'\right)\right) \\ &- \left(\mathbb{E}\left[\max_{i' \in \mathcal{D}_R(d)} \left\{v\left(i', x'\right) + \eta'_{i'}\right\}\right] - v\left(d', x'\right)\right).\end{aligned} \tag{4}$$

Arcidiacono and Miller (2011) show that both terms of the difference on the right-hand side of equation (4) are a function of the (genuine) choice probabilities and thus are identified.

$\boldsymbol{\Theta}(d)$ reflects the additional expected future surplus from not being constrained in one's choice. In the rank condition, this surplus is weighted by the probability of reaching the respective state x and the difference in the constraining probabilities between states A and B.

Note that assumption 7 allows that $\pi(r, x^A) = \pi(r, x^B)$. In this case, the rank condition is always fulfilled. By assumption, the difference between the constraining probabilities for alternative ℓ must be different from zero. In addition, the term $\boldsymbol{q}\left(\ell, x^A\right)^{\top} \boldsymbol{\Theta}(\ell)$ must be greater than zero; otherwise, individuals are only constrained in states that they do not reach or have no value for the excluded choices. Under assumption 3 and the finite support of $\mathcal{X}$, $\boldsymbol{q}\left(\ell, x^A\right)^\top \boldsymbol{\Theta}\left(\ell\right)$ is strictly greater than zero. If $\pi(r, x^A) \neq \pi(r, x^B)$, the rank condition is violated if both choices l and r have the same expected surplus of not being constrained multiplied by the difference in constraining probabilities.

Theorem 1. *Under assumptions 1-8 (let $x^A, x^B \in \mathcal{X}$ and $\ell, r \in \mathcal{D}_G$ be the states and alternatives that fulfil assumption 7), the exponential discount factor β is point identified by*

$$\beta = \frac{\ln\left(\frac{\Pr\left(\ell|\mathcal{D}_G, x^A\right)}{\Pr\left(\ell|\mathcal{D}_G, x^B\right)}\right) - \ln\left(\frac{\Pr\left(r|\mathcal{D}_G, x^A\right)}{\Pr\left(r|\mathcal{D}_G, x^B\right)}\right)}{\left(\pi\left(\ell, x^B\right) - \pi\left(\ell, x^A\right)\right)\boldsymbol{q}\left(\ell, x^A\right)^\top \boldsymbol{\Theta}\left(\ell\right) - \left(\pi\left(r, x^B\right) - \pi\left(r, x^A\right)\right)\boldsymbol{q}\left(r, x^A\right)^\top \boldsymbol{\Theta}\left(r\right)}. \tag{5}$$

Proof. Fix $d' \in \mathcal{D}_G$. Subtract the value function $v\left(d', x'\right)$ from both terms within the square brackets in (2), and add it once to neutralize the subtraction to obtain

$$\begin{aligned} v\left(d, x\right) = u\left(d, x\right) + \beta \sum_{x' \in \mathcal{X}} \Big[&\left(1 - \pi\left(d, x\right)\right) m\left(\mathcal{D}_G, d', x'\right) \\ &+ \pi\left(d, x\right) m\left(\mathcal{D}_R\left(d\right), d', x'\right) + v\left(d', x'\right) \Big] q\left(x' \middle| d, x\right), \end{aligned} \tag{6}$$

where $m\left(\mathcal{D}, d', x'\right) = \ln\left(\sum_{i' \in \mathcal{D}} \exp\left(v\left(i', x'\right)\right)\left(\exp\left(v\left(d', x'\right)\right)\right)^{-1}\right)$ for a given choice set $\mathcal{D} \in \left\{\mathcal{D}_G, \mathcal{D}_R\left(d\right); d \in \mathcal{D}_G\right\}$.

Let $\boldsymbol{m}\left(\mathcal{D}, d'\right)$, $\boldsymbol{q}\left(d, x\right)$, and $\boldsymbol{v}\left(d'\right)$ denote vectors of size $J \times 1$, of which the j-th element is $m\left(\mathcal{D}, d', x_j\right)$, $q\left(x_j \middle| d, x\right)$ and $v\left(d', x_j\right)$, respectively. Using this notation, (6) can be expressed as

$$\begin{aligned} v\left(d, x\right) = u\left(d, x\right) + \beta \Big[&\left(1 - \pi\left(d, x\right)\right) \boldsymbol{q}\left(d, x\right)^\top \boldsymbol{m}\left(\mathcal{D}_G, d'\right) \\ &+ \pi\left(d, x\right) \boldsymbol{q}\left(d, x\right)^\top \boldsymbol{m}\left(\mathcal{D}_R\left(d\right), d'\right) + \boldsymbol{q}\left(d, x\right)^\top \boldsymbol{v}\left(d'\right) \Big]. \end{aligned} \tag{7}$$

Hotz and Miller (1993) show that for a given state x, the difference between the value functions of two alternatives can be identified by some function of their choice probabilities. Under assumption 3, the difference between the value functions of two alternatives $\ell, r \in \mathcal{D}_G$

is determined by the difference in the logarithms of their genuine choice probabilities:

$$\ln\left(\Pr\left(\ell\left|\mathcal{D}_G, x\right.\right)\right) - \ln\left(\Pr\left(r\left|\mathcal{D}_G, x\right.\right)\right) = v\left(\ell, x\right) - v\left(r, x\right). \tag{8}$$

By combining (7) and (8), the following can be derived

$$\begin{aligned}
\ln\left(\Pr\left(\ell\left|\mathcal{D}_G, x\right.\right)\right) - \ln\left(\Pr\left(r\left|\mathcal{D}_G, x\right.\right)\right) &= u\left(\ell, x\right) - u\left(r, x\right) \\
&+ \beta\Big[\left(1 - \pi\left(\ell, x\right)\right)\boldsymbol{q}\left(\ell, x\right)^\top \boldsymbol{m}\left(\mathcal{D}_G, d'\right) + \pi\left(\ell, x\right)\boldsymbol{q}\left(\ell, x\right)^\top \boldsymbol{m}\left(\mathcal{D}_R\left(\ell\right), d'\right) \\
&\quad - \left(1 - \pi\left(r, x\right)\right)\boldsymbol{q}\left(r, x\right)^\top \boldsymbol{m}\left(\mathcal{D}_G, d'\right) - \pi\left(r, x\right)\boldsymbol{q}\left(r, x\right)^\top \boldsymbol{m}\left(\mathcal{D}_R\left(r\right), d'\right) \\
&\quad + \left(\boldsymbol{q}\left(\ell, x\right) - \boldsymbol{q}\left(r, x\right)\right)^\top \boldsymbol{v}\left(d'\right)\Big].
\end{aligned} \tag{9}$$

Most elements in (9) are identified. The unknown elements are β, the difference in instantaneous utilities between the two alternatives ℓ and r, and the value functions of alternative d'.[10]

Subtracting (9) using $x = x^B$ from the same equation using $x = x^A$ results in

$$\begin{aligned}
&\ln\left(\frac{\Pr\left(\ell\left|\mathcal{D}_G, x^A\right.\right)}{\Pr\left(\ell\left|\mathcal{D}_G, x^B\right.\right)}\right) - \ln\left(\frac{\Pr\left(r\left|\mathcal{D}_G, x^A\right.\right)}{\Pr\left(r\left|\mathcal{D}_G, x^B\right.\right)}\right) = \\
&\beta\Big[\left(\pi\left(\ell, x^B\right) - \pi\left(\ell, x^A\right)\right)\boldsymbol{q}\left(\ell, x^A\right)^\top \boldsymbol{\Theta}\left(\ell\right) - \left(\pi\left(r, x^B\right) - \pi\left(r, x^A\right)\right)\boldsymbol{q}\left(r, x^A\right)^\top \boldsymbol{\Theta}\left(r\right)\Big].
\end{aligned} \tag{10}$$

Using assumption 8, both sides of equation (10) can be divided by the right-hand-side term in square brackets. ■

Corollary 1.1. *Under assumptions 1-8, two states $x^A, x^B \in \mathcal{X}$ and two alternatives $\ell, r \in \mathcal{D}_G$ that fulfil assumption 7 and $\pi(r, x^A) = \pi(r, x^B)$, the exponential discount factor β is point identified by*

$$\beta = \frac{\ln\left(\Pr\left(\ell\left|\mathcal{D}_G, x^A\right.\right)\right) - \ln\left(\Pr\left(\ell\left|\mathcal{D}_G, x^B\right.\right)\right)}{\left(\pi\left(\ell, x^B\right) - \pi\left(\ell, x^A\right)\right)\boldsymbol{q}\left(\ell, x^A\right)^\top \boldsymbol{\Theta}\left(\ell\right)}. \tag{11}$$

Proof. Given $\pi(r, x^A) = \pi(r, x^B)$ and assumption 7, the expected future is equal for both

10. If $d' \notin \mathcal{D}$, then $\boldsymbol{m}\left(\mathcal{D}, d'\right)$ is not identified for all distributions of the preference shock. The identifying equation (5), however, depends on $\boldsymbol{\Theta}\left(d\right), d \in \ell, r$, for which d' can be chosen such that $d' \in \mathcal{D}_R\left(d\right)$.

states x^A and x^B when choosing alternative r. Thus, $\Pr\left(r \left| \mathcal{D}_G, x^A\right.\right) = \Pr\left(r \left| \mathcal{D}_G, x^B\right.\right)$. ∎

Equation (11) has a clear economic interpretation. The rank condition guarantees that the expected futures for states x^A and x^B have different values for the decision maker. If the genuine choice probabilities for both states are identical, the decision maker ignores the future consequences and is fully myopic. This aligns with the mathematical result that β is zero in such cases.

For decision makers who are not fully myopic, it is important to check whether β is positive. Assume without loss of generality that $\pi(\ell, x^A) < \pi(\ell, x^B)$. In this case, the denominator of equation (11) is positive. The numerator is only positive if $\Pr\left(\ell \left| \mathcal{D}_G, x^A\right.\right) > \Pr\left(\ell \left| \mathcal{D}_G, x^B\right.\right)$, that is, only if the decision makers choose alternative ℓ more often after state x^A than after state x^B. The model of decision making presented here guarantees that this is the case, as the only difference between the two states is that decision makers are more often constrained in their choices when in state x^B. Thus, a $\beta < 0$ is an indicator of model misspecification.

The size of β depends on three factors: the difference in the genuine choice probabilities, the difference between the probability of being constrained, and a term that is related to the additional expected value of not being constrained. All three terms have the expected impact. Holding the other two constant, β grows with the difference between the choice probabilities. Intuitively, the greater the extent to which choices differ for a given difference in the future expected value is, the more individuals take their expected future into account.

Given an observed difference in choice probabilities, the larger the difference in expected values between being constrained and not, the smaller β is. The result is equivalent for the difference in the constraining probabilities. Intuitively, the greater the extent to which the expected future varies for a given difference in choice probabilities, the more myopic individuals are.

The discussion about the economic meaning of equation (11) applies also to equation (5). For the latter equation, similar arguments regarding choice r can be made. Note that for r, the three discussed factors enter with the opposite sign. The rank condition guarantees that for choices ℓ and r, the weighted surplus of not being constrained in the future differs.

Thus, if the numerator of equation (5) equals zero, it again aligns with decision makers being myopic.

4 Identification of instantaneous utilities

This section extends Corollary 2 of Magnac and Thesmar (2002) to models with choice constraints. Corollary 2.1 states a sufficient condition for the identification of instantaneous utilities for a stationary dynamic discrete choice model with an infinite horizon and at least one strictly positive constraining probability.

To do so, one further assumption is necessary.

Assumption 9. (Repetitive choice possibility). *There exists a $d \in \mathcal{D}_G$, such that $d' \in \mathcal{D}_R(d)$ with $d' = d$.*

This assumption is likely to be fulfilled in many settings. For example, in the context of labor supply, a nonemployed individual can always choose nonemployment. In the context of consumer decisions, a consumer can always choose not to buy anything. In the context of firm competition, a firm can always choose to not operate in a market. For frameworks that do not feature an alternative that is included in its own constrained choice set, researchers can assume a preference shock distribution that fulfils the assumption of the independence of irrelevant alternatives (IIA).

In the following corollary, conditions 1-3 are identical to those of Corollary 2 of Magnac and Thesmar (2002); only condition 4 is newly added.

Corollary 2.1. *Let assumptions 1-6 be fulfilled. Given that the data $\{\Pr(d|\mathcal{D}_G, x), q(\cdot|d, x), \pi(d, x); (d, x) \in \mathcal{D}_G \times \mathcal{X}\}$ are available for at least two consecutive periods, all instantaneous utilities can be recovered if the following assumptions are fulfilled:*

(1) The distribution of preference shocks is known.

(2) The instantaneous utility of one alternative is normalized.

(3) The discount factor β is known.

(4) Assumption 9 is fulfilled.

Proof. Let $e \in \mathcal{D}_G$ with $e \in \mathcal{D}_R(e)$. Let $\mathbf{\Pi}(d)$ be the diagonal matrix with i-th diagonal entry $\pi(d, x_i)$. Let $\boldsymbol{Q}(d)$ denote a matrix with the $\{i,j\}$-th element being $q(x_j | d, x_i)$. Finally, let $\boldsymbol{I}_J$ denote an identity matrix of size J. Given this notation, (6) can be expressed for the whole state space as

$$\boldsymbol{v}(e) = \boldsymbol{u}(e) + \beta\Big[(\boldsymbol{I}_J - \mathbf{\Pi}(e))\,\boldsymbol{Q}(e)\,\boldsymbol{m}(\mathcal{D}_G, e) + \mathbf{\Pi}(e)\,\boldsymbol{Q}(e)\,\boldsymbol{m}(\mathcal{D}_R(e), e) + \boldsymbol{Q}(e)\,\boldsymbol{v}(e)\Big].$$

Minor manipulation results in

$$\begin{aligned}\boldsymbol{v}(e) &= [\boldsymbol{I}_J - \beta\boldsymbol{Q}(e)]^{-1}\,\boldsymbol{u}(e)\\ &\quad + [\boldsymbol{I}_J - \beta\boldsymbol{Q}(e)]^{-1}\,\beta\Big[(\boldsymbol{I}_J - \mathbf{\Pi}(e))\,\boldsymbol{Q}(e)\,\boldsymbol{m}(\mathcal{D}_G, e) + \mathbf{\Pi}(e)\,\boldsymbol{Q}(e)\,\boldsymbol{m}(\mathcal{D}_R(e), e)\Big].\end{aligned}$$

Note that $\boldsymbol{m}(\mathcal{D}, e)$, with $\mathcal{D} \in \{\mathcal{D}_G, \mathcal{D}_R(d); d \in \mathcal{D}_G\}$ can be directly identified from the data using the genuine choice probabilities as $e \in \mathcal{D}_R(e)$. Under assumption 3, the j-th element of $\boldsymbol{m}(\mathcal{D}, e)$ is identified by

$$m(\mathcal{D}, e, x_j) = \ln\left(\frac{\sum_{i \in \mathcal{D}} \Pr(i | \mathcal{D}_G, x_j)}{\Pr(e | \mathcal{D}_G, x_j)}\right).$$

Knowing β and normalizing $\boldsymbol{u}(e) = \mathbf{0}$ uniquely determines $\boldsymbol{v}(e)$. Having $\boldsymbol{v}(e)$ recovered, all other value functions are determined by a combination of the genuine choice probabilities and $\boldsymbol{v}(e)$. With assumption 3,

$$v(d, x) = \ln\left(\frac{\Pr(d | \mathcal{D}_G, x)}{\Pr(e | \mathcal{D}_G, x)}\right) - v(e, x)$$

can be derived. Finally, utilities are uniquely determined by

$$\begin{aligned}\boldsymbol{u}(d) &= \boldsymbol{v}(d) - \beta\Big[(\boldsymbol{I}_J - \mathbf{\Pi}(d))\,\boldsymbol{Q}(d)\,\boldsymbol{m}(\mathcal{D}_G, e)\\ &\qquad + \mathbf{\Pi}(d)\,\boldsymbol{Q}(d)\,\boldsymbol{m}(\mathcal{D}_R(d), e) + \boldsymbol{Q}(d)\,\boldsymbol{v}(e)\Big],\ \forall d \in \mathcal{D}_G \backslash \{e\}.\end{aligned}$$

■

5 Identifying constraining probabilities

In some cases, constraining probabilities are unknown and cannot be estimated from other data sources. For instance, in the context of labor supply, job offers are rarely observed. As a direct result, the researcher can also not observe from which choice set a decision maker has chosen. In such cases, the researcher has to disentangle choice constraints and preferences while only observing composite choice probabilities. The following presents three different sets of additional assumptions; each alone is sufficient for the identification of all constraining probabilities.

The assumptions of Section 2 are assumed to be fulfilled, as the derivation relies especially on assumptions 1, 2, 4, and 5. Furthermore, the researcher knows all *composite* choice probabilities.

Assumption 10. (No constraints after one choice). *There exists a $d \in \mathcal{D}_G$, such that $\pi(d,x) = 0$, $\forall x \in \mathcal{X}$.*

This assumption is fulfilled, for example, in a simple labor supply model with the two choices: *unemployment* and *employment*. Unemployed individuals require a job offer to transition to employment. Employed individuals, however, can freely choose between the two choices.

Proposition 1. *Under assumption 10, all constraining probabilities are identified.*

Proof. See Appendix A.1.

If assumption 10 is not fulfilled, constraining probabilities can only be recovered from the composite choice probabilities if at least one constraining probability is known. Two cases are discussed.

Assumption 11. (Singleton as a constrained set). *There exists a $d \in \mathcal{D}_G$, such that $\mathcal{D}_R(d)$ is a singleton. Further, $\pi(d,x)$ is known $\forall x \in \mathcal{X}$, with $\pi(d,x) < 1$ $\forall x \in \mathcal{X}$.*

Assumption 11 might be fulfilled for models sketched after assumption 10 but including involuntary lay-offs. Thus, choosing employment might also lead to a constrained choice set in the next period. If an individual is laid off, he or she is forced to choose nonemployment. Lay-off probabilities can sometimes be recovered from data on plant closures or surveys.[11] Thus, the researcher only needs to identify the job offer probabilities from the composite choice probabilities.

Proposition 2. *Under assumption 11, all constraining probabilities are identified.*

Proof. See Appendix A.2.

The last sufficient condition discussed here regards alternatives that are excluded in at least two constrained choice sets.

Assumption 12. (Multiple excluded alternatives). *There does not exist a $d' \in \mathcal{D}_G$ such that $\exists! d \in \mathcal{D}_G$ with $d' \in \mathcal{D}_R(d)$. Furthermore, $\exists d \in \mathcal{D}_G$, such that $\pi(d, x)$ is known $\forall x \in \mathcal{X}$.*

An example for this condition is a labor supply model with involuntary separations and the three alternatives *out of the labor force*, *unemployed* and *employed*. The constrained choice sets when not receiving a job offer or being laid off consist of the two alternatives *unemployment* and *out of the labor force*. Again, the probability of an involuntary separation has to be known.

Proposition 3. *Under assumption 12, all constraining probabilities are identified.*

Proof. See Appendix A.3.

6 Identification of genuine choice probabilities

In demand models that features product availability, it might often be possible for a researcher to observe the set from which a customer chooses. In other cases, the researcher is not able

11. For example, the German Socio-Economic Panel specifically asks interviewees if they were involuntary laid off if they report a job transition.

to do so. For instance, in a labor supply model with job offer probabilities, it seems unlikely that the researcher can observe whether an individual was able to choose employment. In such a case, the researcher needs to derive the genuine choice probabilities from the composite choice probabilities.

Throughout this section, it is assumed that the data $\{\text{KPr}(d'|d,x,x'), \pi(d,x); d', d \in \mathcal{D}_G, x', x \in \mathcal{X}\}$ are given for at least two consecutive periods. Note that Section 5 discusses the identification of the constraining probabilities without assuming that the genuine choice probabilities are known. Furthermore, the model in Section 2 is assumed; in particular assumptions 1, 2, 4 and 5 are assumed to be fulfilled.

The following lemma simplifies the discussion:

Lemma 1. *Let $d, d' \in \mathcal{D}_G$, such that $d' \in \mathcal{D}_R(d)$. Furthermore, $\exists x \in \mathcal{X}$, such that $\pi(d,x) > 0$. Then, for $x' \in \mathcal{X}$, $\Pr(d'|\mathcal{D}_G, x')$ identifies $\Pr(d'|\mathcal{D}_R(d), x')$ $\forall d' \in \mathcal{D}_R(d)$.*

Proof. See Appendix B.1.

Given Lemma 1, the focus is on the identification of the genuine choice probability when choosing from the general choice set $\mathcal{D}_G$. Note that if for a given $d \in \mathcal{D}_G$ $\nexists x \in \mathcal{X}$, such that $\pi(d,x) > 0$, then $\mathcal{D}_R(d)$ cannot be reached, and all $\Pr(\cdot|\mathcal{D}_R(d), \cdot)$ are irrelevant.

It is first shown that under assumption 10, the genuine choice probabilities can also be identified.

Proposition 4. *Under assumption 10, all genuine choice probabilities are identified.*

Proof. See Appendix B.2.

Note that assumption 10 can be slightly relaxed for the genuine choice probabilities. It is sufficient that for every state $x \in \mathcal{X}$, $\exists d \in \mathcal{D}_G$, such that $\pi(d,x) = 0$.

Four further cases are discussed.

Assumption 13. *Let $e' \in \mathcal{D}_G$. For each alternative, $d' \in \mathcal{D}_G \backslash \{e'\}$ $\exists d \in \mathcal{D}_G$, such that $d' \notin \mathcal{D}_R(d)$.*

Assumption 13 is a less restrictive version of assumption 11. Therefore, it includes models that feature a constrained choice set that is a singleton.

Proposition 5. *Under assumption 13, all genuine choice probabilities are identified.*

Proof. See Appendix B.3.

The final two assumptions relate to assumption 12. First, a special case is discussed that allows for the identification of the genuine choice probabilities. Thereafter, a more general case is discussed.

Assumption 14. *There exists a $d, e \in \mathcal{D}_G$, $d \neq e$, such that $\mathcal{D}_R(d) = \mathcal{D}_R(e)$, with $\pi(d, x) \neq \pi(e, x) \,\forall x \in \mathcal{X}$. Furthermore, $\pi(d, x), \pi(e, x) > 0 \forall x \in \mathcal{X}$.*

Proposition 6. *Under assumption 14, all genuine choice probabilities are identified.*

Proof. See Appendix B.4.

The last assumption concerns alternatives that are common across all constrained choice sets. It shows that genuine choice probabilities are identified as long as there is sufficient variation in the constraining probabilities.

Assumption 15. *For all $d', e' \in \mathcal{D}_G, d' \neq e'$ with $d', e' \in \mathcal{D}_R(d) \,\forall d \in \mathcal{D}_G$, $\pi(d, x) \neq \pi(e, x)$.*

Proposition 7. *Under assumption 15, all genuine choice probabilities are identified.*

Proof. See Appendix B.5.

7 Conclusion

This paper presents a new exclusion restriction to identify the exponential discount factor in dynamic discrete choice models. It relies on the observation that decision makers are often constrained in their choice. Variation in the probability of being constrained can identify agents' time preferences. The discussed model framework also includes models featuring deterministic variation in decision makers' choice sets.

The presented exclusion restriction requires variation that exclusively affects different constraining probabilities. This variation is not allowed to cause differences in instantaneous utilities or transition probabilities. Under these conditions, the exponential discount factor is point identified.

External factors such as the opposite side of a market or regulators often constrain choices. As such external factors might exclusively impact agents' possibilities to choose from all alternatives and nothing else, the presented exclusion restriction might be fulfilled in numerous applications. Examples that satisfy the exclusion restriction are sketched, including models of labor supply, product demand, and mergers and acquisitions. In the labor supply context, the exclusion restriction is fulfilled if there is exogenous variation in job offer probabilities, for example, due to labor demand shocks. In the context of product demand, variation in the availability of products can be used to identify time preferences. For models describing mergers and acquisitions, variations in regulations can be exploited.

The derived identification equation can be economically interpreted. It reflects the intuition that the greater the observed reaction to a given difference in future expected values is, the greater the exponential discount factor.

It is also shown how instantaneous utilities can be recovered in the discussed framework, provided that the exponential discount factor is known. To identify these utilities, a result of Magnac and Thesmar (2002) is adjusted to fit the presented framework. Although the identification of the exponential discount factor does not exploit the stationarity or normalization of the utility of one alternative, the identification of the instantaneous utilities does.

In some contexts, the researcher might be unable to observe whether decision makers are actually constrained when making their choice. For such cases, the observed choice probabilities are a combination of the probability of being constrained and actual preferences. Several sufficient conditions are discussed to identify the probabilities of being constrained and the genuine choice probabilities based on preferences.

The paper provides grounds for future research. For example, it might be possible to include more than one constrained choice set per alternative. Identification seems possible in these cases, but the recovery of the constraining and genuine choice probabilities is more difficult. Future research might also derive conditions to identify the parameters of hyperbolic discounting.

Appendix A: Proofs of Section 5

All proofs in this section are conditioned on $(x, x') \in \mathcal{X} \times \mathcal{X}$ such that $q(x' | d, x) > 0$ for the respective $d \in \mathcal{D}_G$.

Appendix A.1 Proof of Proposition 1

Proof. Let $d, e, d' \in \mathcal{D}_G$, such that $d \neq e$, $d' \notin \mathcal{D}_R(e)$ and $\pi(d, x) = 0$, $\pi(e, x) > 0\ \forall x \in \mathcal{X}$.

Thus, $\mathrm{KPr}(d' | d, x, x') = \Pr(d' | \mathcal{D}_G, x')$ and $\mathrm{KPr}(d' | e, x, x') = (1 - \pi(e, x)) \Pr(d' | \mathcal{D}_G, x')$. Dividing the two equations and rearranging leads to

$$\pi(e, x) = 1 - \frac{\mathrm{KPr}(d' | e, x, x')}{\mathrm{KPr}(d' | d, x, x')},$$

identifying $\pi(e, x)\ \ \forall e \in \mathcal{D}_G, e \neq d$. ■

Appendix A.2 Proof of Proposition 2

Proof. Choose $d \in \mathcal{D}_G$, such that $\mathcal{D}_R(d) = \{e'\}$, $e' \in \mathcal{D}_G$. Thus,

$$\mathrm{KPr}\left(d' \middle| d, x, x'\right) = (1 - \pi(d, x)) \Pr\left(d' \middle| \mathcal{D}_G, x'\right)\ \ \forall d' \in \mathcal{D}_G, d' \neq e'.$$

Knowing $\pi(d, x)$ makes it possible to recover $\Pr(d' | \mathcal{D}_G, x')\ \forall d' \neq e'$. $\Pr(e' | \mathcal{D}_G, x')$ is given by

$$\Pr\left(e' \middle| \mathcal{D}_G, x'\right) = 1 - \sum_{d' \notin \mathcal{D}_R(d)} \Pr\left(d' \middle| \mathcal{D}_G, x'\right).$$

Having all genuine choice probabilities identified, the unknown constraining probabilities are identified by:

$$\pi(f, x) = 1 - \frac{\mathrm{KPr}(d' | f, x, x')}{\Pr(d' | \mathcal{D}_G, x')}\ \ \forall f \in \mathcal{D}_G,\ \text{with } d' \notin \mathcal{D}_R(f).$$

■

Appendix A.3 Proof of Proposition 3

Proof. The composite choice probability of alternative $d' \notin \mathcal{D}_R(d)$ is

$$\mathrm{KPr}\left(d' \left| d, x, x' \right.\right) = (1 - \pi(d, x)) \Pr\left(d' \left| \mathcal{D}_G, x' \right.\right).$$

As long as $\exists e \in \mathcal{D}_G,\ e \neq d$, such that $d' \neq \mathcal{D}_R(e)$, it is possible to divide the composite choice probabilities for the two choices e and d:

$$\frac{\mathrm{KPr}(d' | d, x, x')}{\mathrm{KPr}(d' | e, x, x')} = \frac{(1 - \pi(d, x))}{(1 - \pi(e, x))}.$$

Since it is assumed that one constraining probability is known, it is possible to derive the constraining probabilities for all choices $d \in \mathcal{D}_G$. ∎

Appendix B: Proofs for Section 6

All proofs in this section are conditioned on $(x, x') \in \mathcal{X} \times \mathcal{X}$ such that $q(x' | d, x) > 0$ for the respective $d \in \mathcal{D}_G$.

Appendix B.1 Proof of Lemma 1

Proof. Given $d \in \mathcal{D}_G$, let $x \in \mathcal{X}$ such that $\pi(d, x) > 0$; the genuine choice probabilities for $d' \in \mathcal{D}_R(d)$ are

$$\Pr\left(d' \left| \mathcal{D}_R(d), x' \right.\right) = \frac{\mathrm{KPr}(d' | d, x, x') - (1 - \pi(d, x)) \Pr(d' | \mathcal{D}_G, x')}{\pi(d, x)}. \tag{12}$$

∎

Appendix B.2 Proof of Proposition 4

Proof. Assume that for alternative $d \in \mathcal{D}_G$, $\pi(d, x) = 0\ \forall x \in \mathcal{X}$. Then, $\Pr(d' | \mathcal{D}_G, x') = \mathrm{KPr}(d' | d, x, x')$. ∎

Appendix B.3 Proof of Proposition 5

Proof. Given $d',\ d \in \mathcal{D}_G$, the genuine choice probability of alternative $d' \notin \mathcal{D}_R(d)$ is given by

$$\Pr\left(d' \middle| \mathcal{D}_R(d), x'\right) = \frac{\mathrm{KPr}\left(d' | d, x, x'\right)}{1 - \pi(d, x)}. \tag{13}$$

For the alternative $d' \in \mathcal{D}_R(d)\ \ \forall d$, the genuine choice probability is given by

$$\Pr\left(d' \middle| \mathcal{D}_G, x'\right) = 1 - \sum_{e' \in \mathcal{D}_G} \Pr\left(e' \middle| \mathcal{D}_G, x'\right).$$

■

Appendix B.4 Proof of Proposition 6

Proof. Define $\mathcal{D}_R(d) = \mathcal{D}_R(e) = \mathcal{D}_E$. For $d' \in \mathcal{D}_E$, the composite choice probabilities are

$$\begin{aligned}
\mathrm{KPr}\left(d' \middle| d, x, x'\right) &= (1 - \pi(d, x)) \Pr\left(d' \middle| \mathcal{D}_G, x, x'\right) + \pi(d, x) \Pr\left(d' \middle| \mathcal{D}_E, x'\right) \\
\mathrm{KPr}\left(d' \middle| e, x, x'\right) &= (1 - \pi(e, x)) \Pr\left(d' \middle| \mathcal{D}_G, x, x'\right) + \pi(e, x) \Pr\left(d' \middle| \mathcal{D}_E, x'\right).
\end{aligned}$$

Given $\pi(d, x) \neq \pi(e, x)$, the two linear equations are independent. The genuine choice probability of $d' \in \mathcal{D}_E$ is determined by

$$\Pr\left(d' \middle| \mathcal{D}_G, x'\right) = \frac{\pi(e, x)\,\mathrm{KPr}\left(d' | d, x, x'\right) - \pi(d, x)\,\mathrm{KPr}\left(d' | e, x, x'\right)}{\pi(e, x) - \pi(d, x)}.$$

The genuine choice probabilities for $d \notin \mathcal{D}_E$ are identified by an equation similar to (13). ■

Appendix B.5 Proof of proposition 7

Proof. Denote the alternatives that are included in all constrained choice sets by $\mathcal{D}_C = \{d_1^c, \ldots, d_{N^c}^c\}$. For an alternative $d \in \mathcal{D}_G$, denote alternatives $d' \in \mathcal{D}_R(d)$ with $d' \notin \mathcal{D}_C$ by

$\{d'^{d}_{1}, \ldots, d'^{d}_{N^d}\}$. The following system of equations exists for each $d \in \mathcal{D}_G$:

$$\mathrm{KPr}\left(d'^{c}_{1}|d, x, x'\right) = (1 - \pi(d, x)) \Pr\left(d'^{c}_{1}|\mathcal{D}_G, x'\right) + \pi(d, x) \Pr\left(d'^{c}_{1}|\mathcal{D}_R(d), x'\right) \tag{14.1}$$

$$\vdots$$

$$\mathrm{KPr}\left(d'^{c}_{N^c-1}|d, x, x'\right) = (1 - \pi(d, x)) \Pr\left(d'^{c}_{N^c-1}|\mathcal{D}_G, x'\right) + \pi(d, x) \Pr\left(d'^{c}_{N^c-1}|\mathcal{D}_R(d), x'\right) \tag{14.N-1}$$

$$\begin{aligned}\mathrm{KPr}\left(d'^{c}_{N^c}|d, x, x'\right) = {} & (1 - \pi(d, x)) \left[1 - \sum_{i=1}^{N^c-1} \Pr\left(d'^{c}_{i}|\mathcal{D}_G, x'\right) - \sum_{e' \notin \mathcal{D}_R(d)} \Pr\left(e'|\mathcal{D}_G, x'\right)\right] \\ & + \pi(d, x) \left[1 - \sum_{i=1}^{N^c-1} \Pr\left(d'^{c}_{i}|\mathcal{D}_R(d), x'\right) - \sum_{i=1}^{N^d} \Pr\left(d'^{d}_{i}|\mathcal{D}_R(d), x'\right)\right].\end{aligned} \tag{15}$$

The system consists of N^c linearly independent equations. Each of equations (14.1) – (14.N-1) features two unknowns: $\Pr\left(d'^{c}_{j}|\mathcal{D}_G\right)$ and $\Pr\left(d'^{c}_{j}|\mathcal{D}_R(d)\right)$. Equation (15) does not include any additional unknowns, as choices not included in $\mathcal{D}_R(d)$ can be directly identified by their respective equivalent of equation (13). In total, the system includes $2(N^c - 1)$ unknowns. An additional system for $f \in \mathcal{D}_G$, $f \neq d$ only adds $N^c - 1$ unknowns because the genuine choice probabilities choosing from $\mathcal{D}_G$ are already included in the system of equations for choice d. Furthermore, each additional system adds N^c independent equations as long as $\pi(f, x) \neq \pi(d, x)$. With K being the number of alternatives in $\mathcal{D}_G$, there exist $(1 + K)(N^c - 1)$ unknowns and $K \times N^c$ independent equations. Thus, there are $(K - N^c + 1)$ more equations than unknowns, allowing one to identify all genuine choice probabilities. ∎

References

Abbring, Jaap H. 2010. "Identification of Dynamic Discrete Choice Models." *Annual Review of Economics* 2 (1): 367–394. ISSN: 1941-1383. https://doi.org/10.1146/annurev.economics.102308.124349.

Abbring, Jaap H., and Øystein Daljord. 2020. "Identifying the Discount Factor in Dynamic Discrete Choice Models." *Quantitative Economics* 11 (2): 471–501. ISSN: 1556-5068. https://doi.org/10.2139/ssrn.3051326. arXiv: 1808.10651.

Aguirregabiria, Victor, and Junichi Suzuki. 2014. "Identification and counterfactuals in dynamic models of market entry and exit." *Quantitative Marketing and Economics* 12 (3): 267–304. ISSN: 1573711X. https://doi.org/10.1007/s11129-014-9147-5.

Arcidiacono, Peter, and Robert A. Miller. 2011. "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity." *Econometrica* 79 (6): 1823–1867. ISSN: 0012-9682. https://doi.org/10.3982/ecta7743.

———. 2019. "Nonstationary Dynamic Models with Finite Dependence." *Quantitative Economics* 10 (3): 853–890. ISSN: 1759-7323. https://doi.org/10.3982/qe626.

———. 2020. "Identifying Dynamic Discrete Choice Models off Short Panels." *Journal of Econometrics* 215 (2): 473–485. ISSN: 18726895. https://doi.org/10.1016/j.jeconom.2018.12.025.

Chen, Le-Yu. 2017. "Identification of Discrete Choice Dynamic Programming Models with Nonparametric Distribution of Unobservables." *Econometric Theory* 33 (3): 551–577.

Conlone, Christopher T., and Julie Holland Mortimer. 2013. "Demand Estimation Under Incomplete Product Availability." *American Economic Journal: Microeconomics* 5 (4): 1–30. ISSN: 19457669. https://doi.org/10.1257/mic.5.4.1.

Fang, Hanming, and Yang Wang. 2015. "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions." *International Economic Review* 56 (2): 565–596.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. "Time Discounting and Preference: A Critical Time Review." *Journal of Economic Literature* 40 (2): 351–401.

Goeree, Michelle Sovinsky. 2008. "Limited Information and Advertising in the U.S. Personal Computer Industry." *Econometrica* 76 (5): 1017–1074. ISSN: 0012-9682. https://doi.org/10.3982/ecta4158.

Haan, Peter, Luke Haywood, and Ulrich Schneider. 2017. "Time Preferences and Female Labor Supply." *Mimeo,* Mimeo.

Harford, Jarrad. 2005. "What drives merger waves?" *Journal of Financial Economics* 77 (3): 529–560. ISSN: 0304405X. https://doi.org/10.1016/j.jfineco.2004.05.004.

Hotz, V. Joseph, and Robert A Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The Review of Economic Studies* 60 (3): 497–529. ISSN: 0034-6527. https://doi.org/10.2307/2298122.

Hu, Yingyao, and Yuya Sasaki. 2018. "Closed-Form Identification of Dynamic Discrete Choice Models with Proxies for Unobserved State Variables." *Econometric Theory* 34 (1): 166–185.

Iskhakov, Fedor, Thomas H. Jørgensen, John Rust, and Bertel Schjerning. 2017. "The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks." *Quantitative Economics* 8 (2): 317–365. ISSN: 17597323. https://doi.org/10.3982/QE643. http://doi.wiley.com/10.3982/QE643.

Kalouptsidi, Myrto, Paul T. Scott, and Eduardo Souza-Rodrigues. 2019. "Identification of Counterfactuals in Dynamic Discrete Choice Models." *Quantitative Economics,* no. Forthcoming.

Komarova, Tatiana, Fabio Sanches, Daniel Silva Junior, and Sorawoot Srisuma. 2018. "Joint analysis of the discount factor and payoff parameters in dynamic discrete choice models." *Quantitative Economics* 9 (3): 1153–1194. ISSN: 1759-7323. https://doi.org/10.3982/qe675.

Lalive, Rafael. 2008. "How do extended benefits affect unemployment duration? A regression discontinuity approach." *Journal of Econometrics* 142, no. 2 (February): 785–806. ISSN: 03044076. https://doi.org/10.1016/j.jeconom.2007.05.013. http://linkinghub.elsevier.com/retrieve/pii/S0304407607001170.

Levy, Matthew, and Pasquale Schiraldi. 2021. "Identification of intertemporal preferences from choice set variation." *Working Paper,* ISSN: 1556-5068. https://doi.org/10.2139/ssrn.3532617.

Magnac, Thierry, and David Thesmar. 2002. "Identifying Dynamic Discrete Decision Processes." *Econometrica* 70 (2): 801–816. https://doi.org/10.1111/1468-0262.00306.

McFadden, Daniel. 1978. "Modeling the Choice of Residential Location." *Spatial Interaction Theory and Planning Models* 1:72–77.

Meier, Stephan, and Charles Sprenger. 2015. "Temporal Stability of Time Preferences." *Review of Economics and Statistics* 97 (2): 273–286. ISSN: 1725-2806. https://doi.org/10.1162/REST.

Norets, Andriy, and Xun Tang. 2014. "Semiparametric Inference in Dynamic Binary Choice Models." *The Review of Economic Studies* 81 (3): 1229–1262. ISSN: 0034-6527. https://doi.org/10.1093/restud/rdt050. http://restud.oxfordjournals.org/lookup/doi/10.1093/restud/rdt050.

Rust, John. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55 (5): 999–1033. https://doi.org/10.2307/1911259.

———. 1994. "Structural estimation of markov decision processes." *Handbook of Econometrics* 4:3081–3143. ISSN: 15734412. https://doi.org/10.1016/S1573-4412(05)80020-0.

Schneider, Ulrich. 2019. *Essays on Time Preferences and Expectations in Dynamic Decision Making.* PhD diss., Freie Universität Berlin. https://doi.org/10.17169/refubium-1894.

Shocker, Allan D., Moshe Ben-Akiva, Bruno Boccara, and Prakash Nedungadi. 1991. "Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models, and Suggestions." *Marketing Letters* 2 (3): 181–197. ISSN: 1573059X. https://doi.org/10.1007/BF02404071.

Srisuma, Sorawoot. 2015. "Identification in Discrete Markov Decision Models." *Econometric Theory* 31 (3): 521–538.