Identification of Dynamic Discrete Choice Models with Hyperbolic Discounting Using a Terminating Action

Chao Wang* Stefan Weiergraeber* Ruli Xiao*

June 27, 2023

Abstract

We study the identification of dynamic discrete choice models with hyperbolic discounting using a terminating action. We provide novel identification results for both sophisticated and naive agents' discount factors and their utilities in a finite horizon framework under the assumption of a stationary flow utility. In contrast to existing identification strategies we do not require to observe the final period for the sophisticated agent. Moreover, we avoid normalizing the flow utility of a reference action for both the sophisticated and the naive agent. We propose two simple estimators and show that they perform well in simulations.

Keywords: hyperbolic discounting, dynamic discrete choice model, identification
JEL Codes: C61, C50, C25

*Department of Economics, Indiana University Bloomington. Email: Wang: cw102@iu.edu; Weiergraeber: sweiergr@iu.edu; Xiao: rulixiao@iu.edu. We thank Yonghong An, Marc Chan, Yoosoon Chang, Aprajit Mahajan, Brian McManus, Christian Michel, and Joon Park for valuable comments and feedback.

1 Introduction

Dynamic discrete choice models (DDCMs) are widely used to analyze agents' intertemporal decisions in a wide range of fields, most notably, industrial organization (technology adoption), health (addiction and postponement of medical procedures), education (decisions to continue or drop out of school), labor (decisions to stay in the labor force or retire), and trade (entry decisions into export markets). A key object of interest in these models is the agents' time preference, i.e., how they trade-off the future against the present. Traditionally, researchers have adopted exponential discounting, which results in an intertemporal substitution rate that is constant over time and predicts time-consistent behavior.

An increasing body of evidence indicates that often agents' behavior is not time-consistent. A popular generalization of exponential discounting is *hyperbolic discounting*, which generalizes exponential discounting to allow agents to experience *present-bias*. Present-bias implies that agents place disproportionately more weight on their utility today than their future, while they discount consistently when comparing any two future periods.[1] Empirically identifying discount factors is inherently difficult, already in the simple exponential discounting setting. Compared to exponential discounting, hyperbolic discounting models introduce additional parameters, which leads to considerable challenges for both identification and estimation.

In this paper, we contribute to the relatively new literature on identification of DDCMs with hyperbolic discounting. We focus on the economically relevant class of models with a finite horizon, a terminating action, and stationary flow utilities.[2] We provide novel identification results for both sophisticated and naive agents' discount factors and their flow utilities. Our identification strategy exploits the recursive structure of DDCMs along with variation in the observed conditional choice probabilities (CCPs) over time. The key contribution of this paper is to show how the presence of a terminating action allows us to obtain identification with fewer data requirements and less restrictive assumptions than what is used in the existing literature.

Most importantly, we do not need to observe the final period to identify the discount factors of the sophisticated agent, and we provide identification results for the naive agent.

[1] Evidence for present-biased behavior often comes from lab experiments documenting *preference reversals*. For example, when subjects are asked about whether they prefer 1 dollar today or 2 dollars tomorrow, most subjects would choose 1 dollar today. However, when the same subjects are asked whether they would like 1 dollar one month from now, or 2 dollars one month plus one day later, many subjects choose the 2 dollar payout.

[2] While not every economic decision problem has a terminating action, they are prevalent in many economically relevant settings, for example, in new technology adoption (De Groote and Verboven, 2019), long-term financial product decisions (Blevins et al., 2020), and decisions to continue education or drop out of school (Eckstein and Wolpin, 1999).

Moreover, we avoid having to normalize the flow utility of a reference action for both the sophisticated and the naive agent, which is typically done in empirical work, even though it can have detrimental implications for counterfactual simulations and policy recommendations.

Our dynamic model setup follows the seminal framework by O'Donoghue and Rabin (1999). In this framework agents in different periods are modelled as different selves that are independent across time periods. Time-inconsistent behavior can arise because of different objectives of the *current self* and the *future selves*. The most common approach to parametrize the present-bias problem is by (quasi-)hyperbolic discounting.[3] Quasi-hyperbolic time preferences are also referred to as $\beta\delta$-preferences because the discount factor t periods in the future is given by $\beta\delta^t$. Throughout this paper, we will refer to δ as the (traditional) *exponential discounting parameter*, and β is the *present-bias parameter*. The exponential discounting model is a special case of quasi-hyperbolic discounting with $\beta = 1$ and present-bias implies that $\beta < 1$.

Agents with present-biased preferences are further categorized as either *sophisticated* or *naive* agents. The key difference is that sophisticated agents are aware that they will be present-biased in the future and take this into account when making a decision today. In contrast, naive agents are aware of their contemporaneous present-bias but they believe that they will not exhibit present-bias in the future and discount using an exponential rate tomorrow, even though in the future they will still be present-biased.

In spite of its theoretical popularity hyperbolic discounting models have only recently be investigated empirically. The main reason for this is that the joint identification of both present-bias parameters and exponential discounting parameters is notoriously difficult in general DDCMs. Even in DDCMs with exponential discounting, the discount factor is generally non-parametrically unidentified, see, for example, Rust (1994) and Magnac and Thesmar (2002).[4] As hyperbolic discounting introduces at least one additional parameter, namely, the present-bias parameter β, the discount factors are still non-parameterically unidentified unless special restrictions are imposed.

Present-bias introduces significant identification challenges because of the conflict between the current self and the future selves. This requires substantial adjustments in how to extract information about the long-run value functions and the model primitives from the equilibrium CCPs. Therefore, the identification of DDCMs with hyperbolic discounting is not a trivial extension of the existing approaches. For a finite horizon setup, Abbring,

[3]To be precise, *hyperbolic discounting* is used only in continuous time problems. Quasi-hyperbolic discounting is an approximation of hyperbolic discounting in discrete time, see, Laibson (1997). To ease notation we use the term *hyperbolic discounting* throughout this paper.

[4]Abbring and Ø. Daljord (2020b) study the (set) identification of the discount factor in DDCMs under exclusion restrictions that affect expected discounted future utilities but not current utility.

O. Daljord, and Iskhakov (2019) provide a set identification result for the discount factors of sophisticated agents, and Mahajan, Michel, and Tarozzi (2020) study the identification of different types of agents (sophisticated, naive, and partially naive) under stronger exclusion restriction. Both papers require data on the final period. This data requirement can be restrictive, because agents often face very long-term decision problems, such as paying a mortgage, or investing in a pension account, for which final period data are often hard to obtain.[5]

Motivated by the literature that studies identification of DDCMs with exponential discounting using a terminating action, in particular, Bajari et al. (2016), we study how the identification of DDCMs with hyperbolic discounting is facilitated by the presence of a *terminating action*.[6] When an agent chooses a terminating action, the decision problem is immediately terminated and there are no more choices to be made in the future. The presence of a terminating action considerably simplifies the mapping between choice probabilities and value functions in this setup. Most importantly, differencing the choice probability contrasts (between an arbitrary action and the terminating action) across two periods results in a simple term that contains the discount factor, because the continuation value associated with the terminating action K is zero and the flow utilities cancel out because of the stationarity assumption. This facilitates expressing the discount factors as functions of the observed choice probabilities in two consecutive periods, which can be exploited for identification.

In the main section of this paper, we show how this logic can be extended to general models of hyperbolic discounting with both sophisticated and naive agents. An important feature of hyperbolic discounting models is that there is a difference between the intertemporal tradeoff between today and the future and the relative tradeoff between any two future periods. This divergence substantially complicates the relationship between the choice probabilities and the value functions established in Arcidiacono and Miller (2011, Lemma 1).

Moreover, additional challenge arise in the case of naive agents, because their presently perceived future behavior and the actual future behavior will not coincide; therefore, one cannot simply use observed actions in the future to recover the current self's beliefs about the future behavior. Consequently, comparing variation in the choice probabilities across two periods is not enough. However, it is possible to achieve identification by differencing functions of the choice probabilities across one additional period, which introduces an additional adjustment term.

[5]Similarly to Abbring, O. Daljord, and Iskhakov (2019), Tsubota (2021) studies identification of the sophisticated agent case in a finite horizon framework without a terminating action. He does not require data on the final period, but requires stronger exclusion restrictions on the flow utility function.

[6]Identification using a terminating action can be interpreted as a special case of exploiting a *finite dependence* property, see, for example, Arcidiacono and Miller (2011).

Our identification strategy formalizes the often-used intuition that variation in CCPs over time is informative about the discounting parameters, if the current utility levels are held constant or are controlled for. The stationarity of the flow utility can be interpreted as a special case of an exclusion restriction. Exclusion restrictions are regularly exploited to identify discount factors in dynamic models. For both the sophisticated and the naive agent case, we show that it is not necessary to impose a normalization on the flow utility of a reference action, which is often done in empirical work, but can be detrimental for counterfactual simulations.

We provide two sets of formal identification result. First, we show that, if the final three periods in the data are observed, both the preferences of the sophisticated and the naive agent are identified without any special data on agents' beliefs. In this case, identification is facilitated by the fact that one can directly recover the flow utility contrasts from behavior in the final period. This allows us to treat the flow utilities as known in periods $T-1$ and $T-2$ and observed behavior in these periods can be used to identify the two discounting parameters, i.e., the long-run discount factor and the present-bias parameter. Moreover, we show that the discount factors of the sophisticated agent can be recovered in closed form using an OLS estimator.

Second, if the final periods are not observed, one can identify the preferences of the sophisticated agent with only four consecutive periods of data ($t-3$, $t-2$, $t-1$, and t). This is because we can express the perceived long-run value function as a function of the discount factors, the flow utility associated with the terminating action, and data on the choice probabilities from three periods. With one additional period of data, one can exploit the fact that the flow utilities backed out using $(t-3, t-2, t-1)$ have to be identical to the ones backed out using $(t-2, t-1, t)$, which provides overidentifying restrictions to pin down the discount factors.[7]

Finally, we use Monte Carlo simulations to show that our estimators work well in simulations and provide step-by-step guidance on how to implement our identification and estimation strategy in practice. We also illustrate the biases introduced by imposing a wrong normalization on the flow utility. These results are in line with recent research on how the normalization of a reference utility impacts counterfactual simulations, see, for example, Norets and Tang (2014), Aguirregabiria and Suzuki (2014), and Kalouptsidi, Scott, and Souza-Rodrigues (2021).

[7]This approach can also be applied, if the final periods are observed in the data, but one is not willing to assume that the continuation value after the final period is zero.

Related literature. This paper is related to two strands of literature. First, our identification strategy builds on a small literature that uses terminating actions to identify DDCMs.[8]

Blevins et al. (2020) study the reverse mortgage (HECM) industry and exploit the presence of multiple terminating actions to identify consumers' discount factor in a model with exponential discounting. Similarly to our strategy, they do not need to impose a normalization on the flow utility and do not require to observe the final period in the data either. However, we study more general time preferences, i.e., hyperbolic discounting of both sophisticated and naive agents.

Bajari et al. (2016) study the subprime mortgage market using a model of exponential discounting. They show how a terminating action allows the researcher to identify the model without imposing a normalization on the flow utility as long as the final period is observed. In contrast to their study, we fully exploit the recursive structure of the dynamic decision problem with a terminating action to show that even when the final period is not observed, we do not have to impose a normalization on the flow utility to identify the model. Even more importantly, we show that their line of argument can be extended to the much richer model setup with hyperbolic discounting and both sophisticated and naive agents. Ø. Daljord, Nekipelov, and Park (2019) provide an extension of Bajari et al. (2016) to general discount functions and general exclusion restrictions beyond stationary flow utilities. However, they require a normalization of the flow utility of a reference action, which can be detrimental for counterfactual simulations. In addition, they do not allow for naive agents and require to observe the final period in the data. We relax both of these restrictions and therefore provide more general identification results.

Second, there is a small but growing literature on the identification of DDCMs with hyperbolic discounting. As one of the earliest empirical papers on hyperbolic discounting, Fang and Wang (2015) proposes an identification condition for partially naive agents in infinite horizon DDCMs. Abbring and Ø. Daljord (2020a) suggest some improvements over Fang and Wang (2015). In a similar spirit, Chan (2017) estimates a DDCM with hyperbolic discounting factor to analyze welfare dependence. Abbring, O. Daljord, and Iskhakov (2019) study the identification of discounting parameters and nonparametric utility functions only for sophisticated agents. They do not assume the presence of a terminating action and exploit

[8]There are several papers who study dynamic decision problems with a terminating action. However, none of them allows for hyperbolic discounting. Kalouptsidi (2014) investigates the bulk shipping market and firm (or ship) exit behavior as a terminating action. De Groote and Verboven (2019) estimate a model of solar panel adoption behavior, in which a consumer exits the market once she adopts the solar panel and thus adopting is a terminating action. Eckstein and Wolpin (1999) develop a model of obtaining education and work decision, in which dropping out of school is the terminating action. Colas, Findeisen, and Sachs (2021) design the optimal college financial aid and explore college entering and dropout decisions. Student dropout is the terminating action.

general exclusion restrictions across time, actions or states. Conceptually our identification strategy is similar to theirs in that we also exploit polynomial equations in the discount factors. However, when applied to their setup, our estimator for the discount factor simplifies to an OLS estimator.

Mahajan, Michel, and Tarozzi (2020) develop a model of time preferences for three types of agents (time-consistent, sophisticated, and naive), as well as the weights of these types in the population. They apply the model to study the demand for insecticide-treated nets in rural India. Similarly to Abbring, O. Daljord, and Iskhakov (2019) they rely on exclusion restrictions to identify time preferences. In particular, they use purposefully collected data shifting agents' beliefs, which are assumed to be excluded from the flow utility. Our identification approach differs from theirs in crucial ways. Most importantly, our identification strategy does not require any special data about agents' beliefs and relies only on data on states and observed choices. Compared with both Abbring and Ø. Daljord (2020a) and Mahajan, Michel, and Tarozzi (2020), who derive identification based on backward induction from the final period, we obtain identification of the discount factors without observing the final period. This is an important distinction, because for many applications, data might only be available for a short panel, which does not cover the final period. Finally, the terminating action allows us to avoid a normalization on the flow utility and we can show identification for the naive agents' time preferences without any data on agents' beliefs.

Heidhues and Strack (2021) provide a strong negative identification result for a specific binary choice model of task completion. In particular, they show that when the flow utility is determined by *iid* draws from an unknown distribution, time preferences are generically unidentified unless information on the agents' continuation values are observed. We consider our identification results complementary to their nonidentification result and show for a more conventional discrete choice setup, in which the unobservable shock is additively separable and drawn from a known distribution, how the presence of a terminating action can be used to identify hyperbolic time preferences together with the agents' flow utilities.

Levy and Schiraldi (2023) study the identification of time preferences using choice set variation in an infinite horizon DDCM. There are several important differences between their paper and ours. Most importantly, in order to identify hyperbolic time preferences they require that the agent has access to both an absorbing choice, that commits her to choosing that action in all future periods, and an additional choice that allows the agent to commit already today to the absorbing choice in several periods from now without restricting her choice in the next period (Levy and Schiraldi, 2023, Theorem 4). Without the additional commitment choice, their approach is only able to identify time-consistent preferences. In contrast, our identification strategy purely relies on the presence of a terminating action to identify hyperbolic time preferences. Therefore, our results are not a special case of the results in Levy and Schiraldi (2023). Moreover, they focus on the hyperbolic time preferences of a sophisticated agent only, while we also provide identification results for the naive agent.

The rest of the paper is organized as follows. Section 2 describes a general DDCM setup with hyperbolic discounting. Section 3 is the key section of the paper and discusses conditions for and derivation of the identification of the model primitives for both the sophisticated and the naive agent. In Section 4, we present several estimators that build on our identification proofs. Section 5 illustrates that our proposed estimators perform well in simulations. Section 6 concludes.

2 General Hyperbolic Discounting: Model Setup

Consider a DDCM with decision periods indexed by $t = 1, \dots, T$, where T is finite. Let u_t represent the individual's utility in period t. The (expected) life-time utility for an individual in period t is given by

$$U_t(u_t, u_{t+1}, \cdots, u_T) \equiv u_t + \beta\delta \sum_{t'=t+1}^{T} \delta^{t'-t-1} \mathbb{E} u_{t'}, \tag{1}$$

where $\beta \in (0, 1]$ captures the individual's present-bias, and $\delta \in (0, 1)$ is the exponential discount rate. This setup nests the exponential discounting framework as a special case, i.e., $\beta = 1$ indicates that the agent does not have present-bias.

The agent chooses her action in every period t to maximize her expected life-time utility. The expectation is taken over the distribution of the future utilities, which are determined by the future state variables (observed and unobserved) and the actions taken by the future selves. Consequently, a forward-looking agent needs to predict how she will behave in the future, which depends on how the future selves discount her future utilities. We distinguish two types of present-biased agents: *sophisticated* and *naive*. A sophisticated agent knows that her future self is also present-biased and maximizes the life time utility characterized by Equation (1). A naive agent, however, believes that her future self is time-consistent and will maximize the life-time utility characterized by Equation (1) with $\beta = 1$ in any future period.

To illustrate the difference between a sophisticated and a naive agent, consider the decision process of a consumer who considers the adoption of a solar PV system: A sophisticated agent is aware that it is difficult to make a large upfront investment today, and that it will also be difficult to make that investment in the future, even though she knows that from a long-term perspective she should make the investment. So although there is a conflict between the current self and the long-term self, the agent is aware of this conflict.

In contrast, a naive agent thinks that it is hard to invest only today. The reason is that she believes that her future self is time-consistent and maximizes her life-time utility, even though the actual future self follows the same present-biased preferences as the current self. Therefore, there exists a conflict between the future self as perceived by the current self and the actual future self in the future. As a consequence, a naive agent tends to repeatedly postpone unpleasant actions.

The dynamic process can be summarized as follows. In each period t, the agent chooses an action k from a choice set $\mathcal{D} = \{1, 2, \ldots, K\}$. Prior to making the choice, the agent observes the state variables $x_t \in \mathcal{X} = \{x_1, \cdots, x_J\}$ and $\epsilon_t = \{\epsilon_{1,t}, \ldots, \epsilon_{K,t}\}$, where x_t are observable to both the agent and the econometrician, and ϵ_t are only observable to the agent. Following the existing literature, we assume that the action-specific private information enters the agent's utility function in an additively separable way.

Assumption 1 *(Additive separability) We assume that the per-period utility has the following additively separable feature:*

$$u_t(x_t, \epsilon_t, k) = u_{k,t}(x_t) + \epsilon_{k,t}. \tag{2}$$

The state variables x_t are assumed to have finite support $\mathcal{X}$ and follow a stationary Markov process controlled by the agent's choice. We denote the state transition by $Q_k(x'|x)$, where $k \in \mathcal{D}$ is the chosen action. The private-information shocks $\epsilon_{k,t}$ are independent of x_t, prior states, and past choices. The shocks are also independent over time, across choices, and have joint distribution G that is absolutely continuous with respect to the Lebesgue measure. These assumptions allow us to factor the transition distribution function for (x_t, ϵ_t) as follows:

Assumption 2 *(Conditional independence) The state transition is assumed to have the following structure:*

$$Q(x_{t+1}, \epsilon_{t+1}|k, x_t, \epsilon_t) = Q_k(x_{t+1}|x_t)G(\epsilon_{t+1}). \tag{3}$$

Before characterizing the agent's optimization decision, we introduce some notation. Let σ^t_τ, with $\tau > t$, denote the period t-self's belief about her period τ-self's behavior. Moreover, we define $\boldsymbol{\sigma}^t_{t+1} \equiv \{\sigma^t_{t+1}, \ldots, \sigma^t_T\}$ as the period t-self's belief about all her future selves' behavior.[9]

[9] It is noteworthy that the beliefs of the t-self and the the $(t+1)$-self regarding the future selves starting from period $t+2$ are identical, i.e., $\boldsymbol{\sigma}^t_{t+2} = \boldsymbol{\sigma}^{t+1}_{t+2}$.

The *current choice-specific value function* in period t is the expected overall utility that the agent receives, which depends on her perception of her future selves' behavior. Specifically, we have

$$\begin{aligned}
& w_{k,t}(x_t;\boldsymbol{\sigma}_{t+1}^t) \\
\equiv\ & u_{k,t}(x_t) + \beta\delta\mathbb{E}_{\epsilon_{t+1},x_{t+1},\ldots\epsilon_T,x_T}\Big[\Big(u_{\sigma_{t+1}^t,t+1}(x_{t+1}) + \epsilon(\sigma_{t+1}^t)\Big) \\
& + \delta\Big(u_{\sigma_{t+2}^t,t+2}(x_{t+2}) + \epsilon(\sigma_{t+2}^t)\Big)\ldots + \delta^{T-t-1}\Big(u_{\sigma_T^t,T}(x_T) + \epsilon(\sigma_T^t)\Big)|k,x_t\Big] \\
=\ & u_{k,t}(x_t) + \beta\delta\mathbb{E}_{x_{t+1}}E_{\epsilon_{t+1}}\Big[\Big(u_{\sigma_{t+1}^t,t+1}(x_{t+1}) + \epsilon(\sigma_{t+1}^t)\Big) \\
& +\delta\ \mathbb{E}_{\epsilon_{t+2},x_{t+2},\ldots\epsilon_T,x_T}\Big(\Big(u_{\sigma_{t+2}^t,t+2}(x_{t+2}) + \epsilon(\sigma_{t+2}^t)\Big)\ldots + \delta^{T-t-2}\Big(u_{\sigma_T^t,T}(x_T) + \epsilon(\sigma_T^t)\Big)|x_{t+1},\sigma_{t+1}^t\Big)|k,x_t\Big] \\
=\ & u_{k,t}(x_t) + \beta\delta\mathbb{E}_{x_{t+1}}\left[v_{t+1}(x_{t+1};\boldsymbol{\sigma}_{t+1}^t)|k,x_t\right], \qquad (4)
\end{aligned}$$

where $v_{t+1}(x_{t+1};\boldsymbol{\sigma}_{t+1}^t)$ is the perceived long-run value function, i.e., the continuation value function that an agent in period t believes to encounter in period $t+1$ if her future selves behave according to strategies $\boldsymbol{\sigma}_{t+1}^t$. It is defined as

$$v_{t+1}(x_{t+1};\boldsymbol{\sigma}_{t+1}^t) \equiv \mathbb{E}_{\epsilon_{t+1}}\{u_{\sigma_{t+1}^t,t+1}(x_{t+1}) + \epsilon(\sigma_{t+1}^t) + \delta\mathbb{E}_{x_{t+2}}\left[v_{t+2}(x_{t+2};\boldsymbol{\sigma}_{t+2}^t)|x_{t+1},\sigma_{t+1}^t\right]\}, \qquad (5)$$

where $v_{t+2}(x_{t+2};\boldsymbol{\sigma}_{t+2}^t)$ is the value function under the perceived future self's strategy profile $\boldsymbol{\sigma}_{t+2}^t \equiv \{\sigma_\tau^t\}_{\tau=t+2}^T$ with the unobserved state variable ϵ_{t+2} integrated out. Note that in the expression above the expected value is multiplied by the discount factor δ instead of $\beta\delta$, because the present-bias parameter β does not directly enter into the intertemporal rate of substitution between any two future periods from the point of view of the present. To summarize, a forward-looking agent with present bias faces a dynamic tradeoff that consists of two components. First, compared to the current utility, the total future utility is discounted disproportionally by factor $\beta\delta$. Second, each period in the future is discounted geometrically by the factor δ.

We focus on *perception perfect strategies* (O'Donoghue and Rabin, 1999), which are strategy profiles $\boldsymbol{\sigma} \equiv \{\sigma_t, \boldsymbol{\sigma}_{t+1}^t, \forall t\}$ such that each σ_t is a best response to her perceived future strategy profile $\boldsymbol{\sigma}_{t+1}^t$, so that

$$\sigma_t(x_t,\epsilon_t) = \arg\max_{k\in\mathcal{D}}\{w_{k,t}(x_t;\boldsymbol{\sigma}_{t+1}^t) + \epsilon_{k,t}\}. \qquad (6)$$

Following the literature on two-step estimation initiated by Hotz and Miller (1993), we define the conditional choice probability (CCP) as the probability that a specific action is chosen given the current state variables. Since the mapping between the decision rule and the CCPs is one-to-one (Hotz and Miller, 1993), we can characterize the agent's optimal decision by equilibrium CCPs instead of the decision rules. In equilibrium, the CCPs are determined by

$$\begin{aligned} p_{k,t}(x_t) &\equiv Pr[\sigma_t(x_t,\epsilon_t)=k] = Pr[w_{k,t}(x_t;\boldsymbol{\sigma}^t_{t+1})+\epsilon_k \geq w_{j,t}(x_t;\boldsymbol{\sigma}^t_{t+1})+\epsilon_j, \quad \forall j\neq k], k\in\mathcal{D}; x\in\mathcal{X} \\ &\equiv \Phi_k(w_{1t},...,w_{K,t}) \quad \forall k\in\mathcal{D};\forall x\in\mathcal{X}, \end{aligned} \tag{7}$$

where the mapping Φ_k depends on the distribution of the shocks. Let $p_t(x_t)$ collect the CCPs in period t of all actions conditional on the state variable x_t.

In the finite horizon setting, which is the focus of this paper, a forward-looking agent solves the model using backward induction, starting from the final period with the utility specified as $w_{k,T}(x_T)=u_{k,T}(x_T)$. Therefore, the equilibrium CCPs are non-stationary and equilibrium existence does not require the flow utility to the stationary.

Sophisticated agent. We first characterize the equilibrium CCPs of the sophisticated agent. Let P_t denote the collection of the equilibrium CCPs for all (observed) states x_t in period t, i.e., $P_t \equiv \{p_t(x_t), x_t\in\mathcal{X}\}$. Moreover, $\mathcal{P}_{t+1}\equiv\{P_{t+1},\ldots,P_T\}$, collects all equilibrium CCPs starting from period $t+1$. The sophisticated agent's perception of the future self's strategies is consistent with the strategy actually chosen by the future self. That is, $\boldsymbol{\sigma}^t_{t+1}$ is consistent with the actual equilibrium CCPs $\mathcal{P}_{t+1}$ that all her future selves adopt. Therefore, we can use $\mathcal{P}_{t+1}$ to represent $\boldsymbol{\sigma}^t_{t+1}$. Similar to Arcidiacono and Miller (2011, Lemma 1) and Abbring, O. Daljord, and Iskhakov (2019, Eq. 2, 9–13), the perceived long-run value function $v_{t+1}(x_{t+1};\mathcal{P}_{t+1})$ for a sophisticated agent can be written as

$$\begin{aligned} & v_{t+1}(x_{t+1};\mathcal{P}_{t+1}) \\ =\ & \mathbb{E}_{\epsilon_{t+1}}\left[u_{t+1}(x_{t+1},P_{t+1})+\epsilon(P_{t+1})+\delta\mathbb{E}_{x_{t+2}}\left[v_{t+2}(x_{t+2};\mathcal{P}_{t+2})|x_{t+1},P_{t+1}\right]\right] \\ =\ & \mathbb{E}_{\epsilon_{t+1}}\left[\max_{k\in\mathcal{D}}\left[w_{k,t+1}(x_{t+1},\mathcal{P}_{t+2})+\epsilon(k_{t+1})\right]+\delta(1-\beta)\mathbb{E}_{x_{t+2}}\left[v_{t+2}(x_{t+2};\mathcal{P}_{t+2})|x_{t+1},P_{t+1}\right]\right] \\ =\ & \mathbb{E}_{\epsilon_{t+1}}\left[\max_{k\in\mathcal{D}}\left[w_{k,t+1}(x_{t+1},\mathcal{P}_{t+2})+\epsilon(k_{t+1})\right]\right]+\delta(1-\beta)\mathbb{E}_{x_{t+2}}\left[v_{t+2}(x_{t+2};\mathcal{P}_{t+2})|x_{t+1},P_{t+1}\right] \\ =\ & \mathbb{E}_{\epsilon_{t+1}}\max_{k\in\mathcal{D}}\left[w_{k,t+1}(x_{t+1},\mathcal{P}_{t+2})+\epsilon(k_{t+1})\right]+\delta(1-\beta)\sum_{x_{t+2}\in\mathcal{X}}v_{t+2}(x_{t+2};\mathcal{P}_{t+2})Q_{P_{t+1}}(x_{t+2}|x_{t+1}) \\ =\ & m_K(p_{t+1}(x_{t+1}))+w_{K,t+1}(x_{t+1};\mathcal{P}_{t+2})+\delta(1-\beta)\sum_k\sum_{x_{t+2}}v_{t+2}(x_{t+2};\mathcal{P}_{t+2})Q_k(x_{t+2}|x_{t+1})p_{k,t+1}(x_{t+1}), \\ =\ & m_K(p_{t+1}(x_{t+1}))+w_{K,t+1}(x_{t+1};\mathcal{P}_{t+2})+\delta(1-\beta)\sum_{x_{t+2}}v_{t+2}(x_{t+2};\mathcal{P}_{t+2})\bar{Q}_{t+1}(x_{t+2}|x_{t+1}), \end{aligned} \tag{8}$$

where $\bar{Q}_{t+1}(x_{t+2}|x_{t+1}) \equiv \sum_k Q_k(x_{t+2}|x_{t+1})p_{k,t+1}(x_{t+1})$ and $m_K(p(x)) = \mathbb{E}_\epsilon max_k [w_k(x) - w_K(x) + \epsilon_k]$ is determined by the distribution of the ϵ-shocks. When ϵ_t follows an extreme value distribution, $m_K(p(x)) = \gamma - log(p_K(x))$, where γ is the Euler constant. For ease of notation, we assume that ϵ_t follows a mean zero type one extreme value distribution so that $m_K(p(x)) = -log(p_K(x))$. The first equality holds by definition of the perceived long-run value function in Equation (5). The second equality follows by using the definition of the current choice-specific value function from Equation (4). The third equality holds because ϵ_{t+1} affects the future value only indirectly through p_{t+1}. The first half of the fourth equality holds by the fact that P_{t+1} is the agent's optimal strategy. The first half of the fifth equality uses the definition of the social surplus function (Manski, McFadden, et al., 1981, Ch.5); the second half is obtained by taking the expectation over the distribution of the shocks, which is equivalent to taking the expectation over the distribution of the optimal actions.

Naive agent. The naive agent's actual behavior in the future diverges from the optimal strategies that the current self perceives today. Therefore, we cannot use the observed future CCPs to represent the current self's belief. Consequently, $\boldsymbol{\sigma}_{t+1}^t$ is inconsistent with the equilibrium CCPs $\mathcal{P}_{t+1}$ in the data. To formalize the decision process of the current self, we first characterize the future self's decision problem as perceived by the current self. We first introduce the *choice-specific value function of the next period self as perceived by the current self*

$$z_{k,t+1}^t(x; \boldsymbol{\sigma}_{t+2}^t) = u_{k,t+1}(x) + \delta \sum_{x_{t+2}} v_{t+2}(x_{t+2}; \boldsymbol{\sigma}_{t+2}^t) Q_k(x_{t+2}|x_{t+1}), \tag{9}$$

where $\boldsymbol{\sigma}_{t+2}^t$ describes the current self's perception of her future self's choice. From the perspective of the naive agent's current self, σ_{t+1}^t is determined by

$$\sigma_{t+1}^t(x_{t+1}, \epsilon_{t+1}) = \arg\max_{k \in \mathcal{D}} \{z_{k,t+1}^t(x_{t+1}; \boldsymbol{\sigma}_{t+2}^t) + \epsilon_{k,t+1}\}. \tag{10}$$

The actual decision in period $t+1$, however, is determined by the *choice-specific value function* of the $(t+1)$-self $w_{k,t+1}(x; \boldsymbol{\sigma}_{t+2}^{t+1})$, which is given by

$$w_{k,t+1}(x; \boldsymbol{\sigma}_{t+2}^{t+1}) = u_{k,t+1}(x) + \beta\delta \sum_{x_{t+2}} v_{t+2}(x_{t+2}; \boldsymbol{\sigma}_{t+2}^{t+1}) Q_k(x_{t+2}|x_{t+1}), \tag{11}$$

and characterizes how the current self evaluates the deterministic component of the payoff from choosing k. In contrast, $z_{k,t+1}^t(x; \boldsymbol{\sigma}_{t+2}^t)$ represents how the current self believes how her next-period self to evaluate the payoff from choosing k. Formally, the key difference between

$w_{k,t+1}(x;\boldsymbol{\sigma}^{t+1}_{t+2})$ and $z^t_{k,t+1}(x;\boldsymbol{\sigma}^t_{t+2})$ is that the continuation value in the former is discounted by $\beta\delta$, while the continuation value in the latter is discounted by δ.

We define the future CCP as perceived by the current self $p^t_{k,t+1}(x_{t+1})$ as

$$p^t_{k,t+1}(x_{t+1}) \quad = \quad Pr[z^t_{k,t+1}(x;\boldsymbol{\sigma}^t_{t+2}) + \epsilon_{k,t+1} \geq z^t_{j,t+1}(x;\boldsymbol{\sigma}^t_{t+2}) + \epsilon_{j,t+1}, \quad \forall j \neq k]. \tag{12}$$

Let P^t_{t+1} denote the collection of the period $t+1$ CCPs as perceived by the period t self. Moreover, $\mathcal{P}^t_{t+1}$ collects the perceived CCPs for all future periods starting from $t+1$. Even though the current self's belief is different from the future self's actual behavior, the current self's beliefs can be expressed as a function of the perceived future self's CCPs, so that $\mathcal{P}^t_{t+1}$ is consistent with $\boldsymbol{\sigma}^t_{t+1}$. We can rewrite the perceived long-run value function $v_{t+1}(x_{t+1};\mathcal{P}^t_{t+1})$ for a naive agent as

$$\begin{aligned} v_{t+1}(x_{t+1};\mathcal{P}^t_{t+1}) &= \mathbb{E}_{\epsilon_{t+1}}\{z_{P^t_{t+1},t+1}(x_{t+1},\mathcal{P}^t_{t+2})\} \\ &= \mathbb{E}_{\epsilon_{t+1}} \max_{k\in\mathcal{D}} \left[z_{k,t+1}(x_{t+1},\mathcal{P}^t_{t+2})\right] \\ &= m_K(p^t_{t+1}(x_{t+1})) + z_{K,t+1}(x_{t+1};\mathcal{P}^t_{t+2}). \end{aligned} \tag{13}$$

Consequently, the equilibrium CCPs for a naive agent are characterized by Equations 7, 9, and 13. Note that the perceived CCPs $\mathcal{P}^t_{t+1}$ are never observed in the data, although they affect the equilibrium CCPs of the current self. This greatly complicates the identification strategy.

3 Identification Results

In this section, we show identification of the hyperbolic discounting parameters jointly with the payoff primitives in a DDCM with a finite horizon. We present identification results for both the sophisticated and the naive agent. In Section 3.1 we consider the case in which the researcher observes the final periods in the data, which is the setting that most of the existing literature studies. In Section 3.2 we extend the identification to the scenario in which the researcher does not observe the information in the final periods.

Throughout the paper, we assume that we know the type of each decision maker, i.e., whether she is time-consistent, sophisticated, or naive. In practice, a decision maker's type might be unknown, which results in observing a mixture of all three types in the data. By exploiting the panel data structure and relatively standard assumptions, such as that the type is time invariant, we can identify and estimate the type-specific CCPs using insights from the measurement error or finite mixture literature, see, for example, Hu (2008). After the type-specific CCPs are identified, the main challenge is to map each set of identified

CCPs to one of the three types. This step typically requires additional assumption on how exactly the types differ from each other, which depends crucially on the specific application. Once we have mapped each type to its equilibrium CCPs, we can follow the identification results developed in this paper using the type-specific CCPs to identify the flow utilities and the discount factors separately for each type.

Throughout, we rely on several assumptions. First of all, our key assumption is that there is some terminating action available to the agent.

Assumption 3 *Action K is a terminating action.*

Intuitively, the presence of the terminating action facilitates the identification similarly to the information in the final period in the sense that once this action chosen, there is no future involved. However, the choice of the terminating action depends on the future payoff, because the agent chooses among all possible options and other choices depend on the future payoffs.

For ease of illustration, we assume that the unobserved shocks follow a type-1 extreme value distribution, so that the equilibrium CCPs have the logit structure:[10]

$$p_{k,t}(x_t) \quad = \quad \frac{w_{k,t}(x_t; \boldsymbol{\sigma}_{t+1}^t)}{\sum_{j \in \mathcal{D}} w_{j,t}(x_t; \boldsymbol{\sigma}_{t+1}^t)}. \tag{14}$$

We can then transform the log odds ratios of the equilibrium CCPs into contrasts of the choice-specific value functions:

$$\phi_{kK,t}(x) \equiv logp_{k,t}(x) - logp_{K,t}(x) = w_{k,t}(x) - w_{K,t}(x), \tag{15}$$

which contain information about the flow utility, the discount factor, and the present-bias parameter. This equation implies that the log odds ratio is determined by the difference in flow utilities associated with the two actions (k and K) and the continuation value integrated over the future uncertainty as captured by the state transition matrix $\boldsymbol{Q_k}$, instead of the difference between the transition matrices $\boldsymbol{Q}_k - \boldsymbol{Q}_K$. This allows us to fully represent the expected utility as a function of the model primitives, in particular, a function of the discount parameters δ, β, the state transition matrix $\boldsymbol{Q}_k$, and the future *ex ante* value function $\boldsymbol{v}_{t+1}$. Note that the transition matrix can be directly identified from the data under the standard assumption that the agent has rational expectations about the future evolution of the state variables.

Our second key identification assumption is that the flow utility is stationary. Specifically, to identify the discount parameters, we first control for the direct impact of the flow utility,

[10] Conceptually our arguments go through for more general distributions of the error term.

$\boldsymbol{u}_{k,t} - \boldsymbol{u}_{K,t}$ in Equation (15) by differencing the log odds ratios in two consecutive time periods. If the agent's flow utility differences (across actions k and K) do not change over time, which is the case if the flow utility is stationary, the flow utility terms in the log odds ratio equation cancel out. Therefore, we make the following assumption.

Assumption 4 *(Stationary flow utility) The flow utility is time-invariant.*

$$u_{k,t}(x) = u_k(x), \quad \forall x \in \mathcal{X}. \tag{16}$$

The stationarity of the flow utility is reasonable in many economic applications, and this assumption is frequently used in the literature, see, for example, Bajari et al. (2016), Blevins et al. (2020), and An, Hu, and Xiao (2021).

Abbring, O. Daljord, and Iskhakov (2019) discuss that identification of the hyperbolic discounting parameters for a sophisticated agent can be achieved by using a suitable normalization of the flow utility and exclusion restrictions, which either can be satisfied by a stationary utility function so we can exploit the variation of CCPs across time, or by an additional state variable that is excluded from the current utility but affects the future value by shifting the state transition so that one can exploit the variation of CCPs for different values of such an exclude state variable.

3.1 Setting 1: Final Periods Observed

We first study the identification of both the sophisticated and the naive agent's preferences in a scenario in which the researcher observes the data in the final three periods, i.e., $\{\boldsymbol{p}_T, \boldsymbol{p}_{T-1}, \boldsymbol{p}_{T-2}\}$ are known. This is the case that has been studied in the existing literature. In the following we show how the presence of a terminating action allows us to derive identification under more general assumptions, in particular, our strategy does not require a normalization of the flow utility.

First of all, in the final period, the optimal decision is the same regardless of the agent's type, because there is no future to discount, if we assume the continuation value is zero. Consequently, independently of whether the agent is time-consistent, sophisticated or naive, we can directly recover her flow utility contrasts using the log odds ratios in the final period. Our starting point is the log odds ratio vector which is obtained by stacking the log odds ratios from Equation (15) for each state value x:

$$\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T) = \boldsymbol{w}_{k,T} - \boldsymbol{w}_{K,T} = \boldsymbol{u}_k - \boldsymbol{u}_K. \tag{17}$$

The stacked choice-specific value function $\boldsymbol{w}_{k,t}$ and flow utility function $\boldsymbol{u}_k$ are defined analogously. Observing the action in the final period greatly simplifies the identification procedure for both sophisticated and naive agents because we can treat the utility contrast as known when exploiting information on log odds ratios in earlier periods. Moreover, with a normalization assumption on the flow utility $\boldsymbol{u}_K$, the flow utility associated with other actions $\boldsymbol{u}_k$ is identified as in the existing literature. However, we show that such a normalization assumption is not necessary when there is a terminating action. We demonstrate in Section 5 that imposing such a normalization on the flow utility can substantially affect counterfactual simulations.

For now, we assume that the utility contrasts are known[11] and exploit the variation in other periods to identify the flow utility separately for each action and the discount factors.

In the penultimate period $T-1$ the current self has to predict the optimal strategy of her T-self.[12] The sophisticated $(T-1)$-self correctly believes that her T-self will discount hyperbolically when she enters period T. The naive $(T-1)$-self wrongly believes that her T-self is time-consistent, even though she will not be once she is in period T. However, since the T-self faces a static decision, it does not matter how she discounts the future. Consequently, the divergence between the naive $(T-1)$-self's perception about her behavior in period T and her actual behavior in T does not matter, i.e., $\boldsymbol{p}_T^{T-1} = \boldsymbol{p}_T$. It is worth noting that this feature is only true for the second to final period. Therefore, we do not have to take a stance about the agent's time preference, and we can write the log odds ratio of the CCPs in period $T-1$ as

$$\begin{aligned} \boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-1}) &= \boldsymbol{u}_k - \boldsymbol{u}_K + \beta\delta \boldsymbol{Q}_k \boldsymbol{v}_T(\boldsymbol{p}_T^{T-1}) \\ &= \boldsymbol{u}_k - \boldsymbol{u}_K + \beta\delta \boldsymbol{Q}_k \boldsymbol{v}_T(\boldsymbol{p}_T) \\ &= \boldsymbol{\phi}_{kK}(\boldsymbol{p}_T) + \beta\delta \boldsymbol{Q}_k \left(-log \boldsymbol{p}_{KT} + \boldsymbol{u}_K\right), \end{aligned} \tag{18}$$

where $\boldsymbol{v}_T(\boldsymbol{p}_T^{T-1})$ is the stacked ex ante value function over x, and $\boldsymbol{Q}_k$ is a $J \times J$ transition matrix with element (i, j) containing $Q_k(x_{t+1} = j | x_t = i)$. The first equality holds by definition and the fact that K is a terminating action. The second equality holds because $\boldsymbol{p}_T^{T-1} = \boldsymbol{p}_T$ for both the sophisticated and the naive agent. The third equality holds by the one-to-one mapping between *ex ante* value function and the choice-specific value function and the corresponding CCPs.

Instead of imposing a normalization condition on the flow utility as in the existing literature, we recover the flow utility function $\boldsymbol{u}_K$ as a function of the two discount factors δ and β, which requires the following rank condition.

[11] Usually the utility contrasts can be nonparametrically recovered from data on the CCPs following the identification results of Magnac and Thesmar (2002).

[12] To simplify notation, we label an agent's self in period t as *t-self*.

Assumption 5 *(Full rank condition) There exists an action k other than the terminating action K, such that the state transition matrix $\boldsymbol{Q}_k$, controlled by action k, has full rank.*

This rank condition imposes some restrictions on the state transition controlled by the non-terminating actions. Note that this rank condition is directly testable, because we can estimate the state transition matrix from the observed data. The technical advantage of considering a terminating action is that such a full rank condition imposes very mild restrictions on the transition matrix of the non-terminating action, namely it only requires full rank of $\boldsymbol{Q}_k$ itself. Without the presence of a terminating action we require full rank of the transition matrix difference $(\boldsymbol{Q_k} - \boldsymbol{Q_K})$, which is trivially rank-deficient because every column of the transition matrices $\boldsymbol{Q_k}$ and $\boldsymbol{Q_K}$ sums to 1, so that the difference of the two sums is zero.

The full rank condition and the data on the final two periods allow us to identify the flow utility associated with the terminating action as a closed-form function of the data and the two discount factors because the utility contrasts $\boldsymbol{u}_k - \boldsymbol{u}_K$ are known. That is,

$$\boldsymbol{u}_K = log(\boldsymbol{p}_{KT}) - \frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T), \tag{19}$$

where $\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T) = \boldsymbol{\phi}_{kK}(\boldsymbol{p}_T) - \boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-1})$. Consequently, the flow utility associated with other actions is also identified as a function of the compound discount factor $\beta\delta$. Mahajan, Michel, and Tarozzi (2020) exploit a normalization assumption on the flow utility, i.e., $\boldsymbol{u}_K$ is known, to identify the product of the two discount factors directly from Equation (18).

So far we have exploited all variation in the last two periods, in which sophisticated and naive agents behave identically so that the identification procedure is the same. However, the identification arguments using the third to final period, $T-2$, differ for sophisticated and naive agents, because of their different perception of their $(T-1)$-self. Consequently, we investigate the two types of agents separately in the following two subsections and show that identification does not require any normalization assumptions.

3.1.1 Sophisticated agents

Given that $\boldsymbol{u}_K$ is already identified as a function of the two discount factors, we further exploit variation of the CCPs in period $T-2$. Specifically, for period $T-2$, the differences between the log odds ratios $\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-2})$ for a sophisticated agent are

$$\begin{aligned}\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-2}) &= \boldsymbol{u}_k - \boldsymbol{u}_K + \beta\delta\boldsymbol{Q}_k\boldsymbol{v}_{T-1}(\boldsymbol{p}_{T-1}) \\ &= \boldsymbol{u}_k - \boldsymbol{u}_K + \beta\delta\boldsymbol{Q}_k\left(-log\boldsymbol{p}_{KT-1} + \boldsymbol{u}_K + \delta(1-\beta)\bar{\boldsymbol{Q}}_{T-1}(-log(\boldsymbol{p}_{KT}) + \boldsymbol{u}_K)\right),\end{aligned} \tag{20}$$

where the weighted state transition matrix $\bar{\boldsymbol{Q}}_{T-1}$ is a $J \times J$ matrix with element (i, j) equal to $\sum_k Q_k(x_T = j|x_{T-1} = i)p_{k,T-1}(x_{T-1} = i)$. Equation (20) is obtained by substituting in the recursive relationship in the perceived long-run value function as specified in Equation (8). Almost all components in Equation (20) are identified from data on the last two periods. Only the two discount factors are unknown.

Comparing the log odds ratios in period $T-1$ and $T-2$, we can see that the term $\delta(1-\beta)$ enters the future value of period $T-2$ but not the future value of $T-1$. Consequently, the difference between the log odds ratios for period $T-1$ and $T-2$ provides variation to identify the two discount factors separately. That is,

$$\begin{aligned}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-1}) &\equiv \boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-1}) - \boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-2}) \\ &= \beta\delta\boldsymbol{Q}_k\left(log(\boldsymbol{p}_{KT-1}) - log(\boldsymbol{p}_{KT}) + \delta(1-\beta)\bar{\boldsymbol{Q}}_{T-1}\frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)\right),\end{aligned} \tag{21}$$

which follows from plugging in Equation (20) and substituting the last term in Equation (20) using Equation (18). Equation (21) provides J equation but only contains the two discount factors as unknowns. Furthermore, this system of equations is linear in both δ and $\delta\beta$:

$$\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-1}) = \begin{bmatrix} A & B \end{bmatrix} \times \begin{bmatrix} \delta\beta \\ \delta \end{bmatrix} \equiv \Omega(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T) \times \begin{bmatrix} \delta\beta \\ \delta \end{bmatrix}, \tag{22}$$

where $A \equiv \boldsymbol{Q}_k\left[log(\boldsymbol{p}_{KT-1}) - log(\boldsymbol{p}_{KT}) - \bar{\boldsymbol{Q}}_{T-1}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)\right]$ and $B \equiv \boldsymbol{Q}_k\bar{\boldsymbol{Q}}_{T-1}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)$. Consequently, matrix $\Omega(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)$ has size $J \times 2$, which can be identified and estimated from the data directly. Under a full rank condition for this coefficient matrix, we can separately identify the two discount factors β and δ.

Assumption 6 *Matrix* $\Omega(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)$ *has full column rank, i.e.,* $rk\left(\Omega(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)\right) = 2$.

This condition is not very restrictive in practice in the sense that it is satisfied as long as there exists one row of the matrix that cannot be expressed as a linear combination of any of the other rows. When the support of the state x is large, i.e., J is large, the choice of the rows becomes large so it is easier to satisfy this condition. However, if this assumption fails, the model is under-identified, meaning that one of the parameters, say δ, can always be expressed as a function of the other parameter, β. We summarize the above discussion as

Proposition 1 *If Assumptions 1 to 6 are satisfied and all agents are sophisticated, then all* *flow utility functions, the exponential discount factor δ, and the present-bias parameter β are identified.*

There are several noteworthy distinctions between our identification strategy and those of the existing literature. First, our identification mainly exploits the variation of the CCPs over time with the assumption of a stationary flow utility. The stationary assumption can be interpreted as one type of exclusion restriction, namely that time t is excluded from the flow utility function. In contrast, the existing literature relies on the presence of an additional state variable that is excluded from the flow utility function. The stationarity of the flow utility might be a restrictive assumption in some cases. However, we only require the stationary for a short period; for the last three periods in Proposition 1 and for four consecutive periods in Proposition 4 below. These assumptions are less restrictive than it may seem because an agent's preferences might not change over parts of the data sample even though over the whole sample her preferences might change. Once we identify the discount factors using the subset of periods with constant flow utility, we can identify the time-specific flow utility in any other period.

Second, we exploit the presence of a terminating action in order to avoid the normalization assumption on the flow utility, which the existing literature usually has to impose. This is the key difference between our identification strategy and both Abbring, O. Daljord, and Iskhakov (2019) and Mahajan, Michel, and Tarozzi (2020), who require a normalization of the flow utility associated with a reference action, that yields either zero or a known utility. For example, Mahajan, Michel, and Tarozzi (2020), assume that $\boldsymbol{u}_K$ is known, and that there exists an state variable that affects the continuation value but is excluded from the flow utility. Imposing a normalization condition can be problematic for counterfactual analyses as discussed in a general setting in Kalouptsidi, Scott, and Souza-Rodrigues (2021). We provide an illustration of the generated bias in our simulations in Section 5.

3.1.2 Naive agents

In this section we provide identification results for the naive agent by further exploiting the information in period $T-2$. For the third to last period $(T-2)$ we can write the log odds ratios as

$$\begin{aligned}\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-2}) &= \boldsymbol{u}_k - \boldsymbol{u}_K + \beta\delta\boldsymbol{Q}_k\boldsymbol{v}_{T-1}(\boldsymbol{p}_{T-1}^{T-2}) \\ &= \boldsymbol{\phi}_{kK}(\boldsymbol{p}_T) + \beta\delta\boldsymbol{Q}_k\boldsymbol{v}_{T-1}(\boldsymbol{p}_{T-1}^{T-2}), \end{aligned} \tag{23}$$

where $\boldsymbol{p}_{T-1}^{T-2}$ denotes the CCPs that the naive $(T-2)$-self believes to follow in $T-1$. The first equality holds because of the stationarity of the flow utility, the second equality holds by plugging in the identified utility contrasts from choices in period T. Note that the perceived CCPs for period $T-1$ by the $(T-2)$-self differ from the actual behavior of the $(T-1)$-self, i.e., $\boldsymbol{p}_{T-1}^{T-2} \neq \boldsymbol{p}_{T-1}$. This is because in the $(T-2)$-self's perception, the $T-1$-self is discounting period T's utility using δ while the actual $(T-1)$-self discounts the period T utility using $\delta\beta$.

However, choices in period $T-2$ reveal information about the $(T-2)$-self's perception of the behavior of the $(T-1)$-self and the T-self. Instead of using the actual actions of future selves to recover the current self's beliefs, we use the current self's choices to recover her perceptions about the future. That is, from Equation (23) and with the full rank condition on the transition matrix, imposed by Assumption 5, we have

$$\boldsymbol{v}_{T-1}(\boldsymbol{p}_{T-1}^{T-2}) = \frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\left(\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-2}) - \boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)\right). \tag{24}$$

From the perspective of the $(T-2)$-self, the $(T-1)$-self is time-consistent, so she takes into account only the exponential discount factor. The choice-specific value function of the $(T-1)$-self as perceived by the naive $(T-2)$-self can be written as[13]

$$\begin{aligned} z_{kT-1}^{T-2} &= \boldsymbol{u}_k + \delta\boldsymbol{Q}_k\boldsymbol{v}_T(\tilde{\boldsymbol{p}}_T) \\ &= \boldsymbol{u}_k + \delta\boldsymbol{Q}_k\boldsymbol{v}_T(\boldsymbol{p}_T) \\ &= \boldsymbol{u}_k - \frac{1}{\beta}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T), \end{aligned} \tag{25}$$

where $\boldsymbol{v}_T(\tilde{\boldsymbol{p}}_T)$ captures how the $(T-2)$-self believes her $(T-1)$-self‘' to think about period T. Note that the $(T-2)$-self knows that the $(T-1)$-self knows that the T-self faces a static decision; therefore, the present-bias does not affect the T-self's decision. The first equality holds by definition, the second equality holds because the T-self faces a static decision. The third equality holds by plugging in the relationship of $\delta\boldsymbol{Q}_k\boldsymbol{v}_T(\boldsymbol{p}_T)$ and the log odds ratio contrasts specified in Equation (18).

So far we can identify the ex ante value function as perceived by the $(T-2)$-self using her actual actions and the $(T-1)$-self's choice-specific value function in $(T-2)$-self's perception. The two components can be connected via the social surplus function (Manski, McFadden, et al., 1981):

$$\begin{aligned} \boldsymbol{v}_{T-1}(\boldsymbol{p}_{T-1}^{T-2}) &= log\sum_k exp(z_{kT-1}^{T-2} - z_{KT-1}^{T-2}) + z_{KT-1}^{T-2} \\ &= log\sum_k exp(z_{kT-1}^{T-2} - \boldsymbol{u}_K) + \boldsymbol{u}_K, \end{aligned} \tag{26}$$

[13]Note that this term is different from w_{kT-1}, which describes the actual choice-specific value function as considered by the current $T-1$-self in $T-1$.

where the second equality holds because K is a terminating action. Consequently, by plugging in the expressions of the ex ante value function identified from Equation (24) and the choice-specific perceived value function identified from Equation (25), we have

$$\begin{aligned} & \frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\left(\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-2}) - \boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)\right) \\ = \; & log\left(\sum_k exp\left(\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T) - \frac{1}{\beta}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)\right)\right) + log(\boldsymbol{p}_{KT}) - \frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T), \end{aligned} \tag{27}$$

which provides J nonlinear equations in only two unknowns, namely the discount factors β and δ.

Assumption 7 *The gradient of the J restrictions has a rank of 2 at the true parameter values, where the restrictions are defined as*

$$\begin{aligned} R^n(\beta,\gamma) \;\equiv\; & \frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\left(\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-2}) - \boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)\right) \\ - \; & \left[log\left(\sum_k exp\left(\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T) - \frac{1}{\beta}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)\right)\right) + log(\boldsymbol{p}_{KT}) - \frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)\right], \end{aligned}$$

and the gradient is $\nabla R^n(\beta,\gamma) \equiv \begin{bmatrix} \frac{\partial R^n}{\partial \beta} \\ \frac{\partial R^n}{\partial \gamma} \end{bmatrix}$*, where* $\gamma = \beta\delta$*.*

Therefore, we can locally identify the two discount factors for naive agents.[14]

Proposition 2 *If Assumptions 1 to 5 and 7 hold, all agents are naive, and if at least the last three period of data are available, then the discount factors (β, δ) are (locally) identified without imposing any further restrictions on the flow utility.*

If Assumption 7 holds, Proposition 2 states that the set of solutions to Equation (27) is composed of locally isolated points. If there are more than one solution to Equation (27), global identification fails and we can only achieve local identification. In practice, however, we can exploit the special structure of the equation to compute the full set of solutions. It is noteworthy that this equation is linear in $\beta\delta$. Therefore, we only need to do a one-dimensional nonlinear search on $0 \leq \beta \leq 1$, and get $\beta\delta$ in the inner loop as a function of β by exploiting the linearity, which is not computationally demanding. Therefore, if we obtain

[14]We follow the definition of *local identification* in Rothenberg (1971). Specifically, (β, δ) are locally identifiable, if there exists an open neighborhood of (β, δ) containing no other (β, δ) which is observationally equivalent. In contrast to global identification, local identification allows for several observationally equivalent parameter values as long as they are isolated from each other.

an identified set, that is composed of discrete isolated points under Assumption 7, we are still able to estimate the full set of points. If Assumption 7 fails, which seems unlikely to happen in practice, we can only identify a continuum of solutions.

It is worth noting that we implicitly assume that the continuation value in the final period is zero. This allows us to identify the flow utility from the actions in the final period directly. To the best of our knowledge, we are the first to study point identification of the naive agent's time-preference in a DDCM framework. The existing literature mainly focuses on identification of the sophisticated agent, with the only exception of Mahajan, Michel, and Tarozzi (2020), which provide set identification results for the naive agent using data on the final three periods and a normalization of the flow utility.

3.2 Setting 2: Final Periods Not Observed

In many applications, the econometrician will not have data up to the final period. For example, in data sets on long-term financial products, such as mortgages, or long-term health studies, one typically does not observe the final period for most individuals. Moreover, it might be restrictive to assume that the continuation value is zero in some settings. In this subsection, we extend our identification strategy to settings in which the final periods are not observed in the data. This framework can also be applied if one observes the final period but is not willing to normalize the continuation value after the final period to zero.

We start by providing identification results for the sophisticated agent. At the end of this section we discuss why providing identification results for the native agent is much more difficult when the final periods are not observed. Our starting point is the log odds ratio vector (stacked over all J states x) in period t from Equation (15):

$$\boldsymbol{\phi}_{kK}(\boldsymbol{p}_t) = \boldsymbol{w}_{k,t} - \boldsymbol{w}_{K,t} = \boldsymbol{u}_{k,t} - \boldsymbol{u}_{K,t} + \beta\delta \boldsymbol{Q}_k \boldsymbol{v}_{t+1}, \tag{28}$$

Given this expression, we can difference the CCPs in two consecutive periods so that

$$\begin{aligned} \Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{t+1}) &\equiv \boldsymbol{\phi}_{kK}(\boldsymbol{p}_{t+1}) - \boldsymbol{\phi}_{kK}(\boldsymbol{p}_t) \\ &= \beta\delta \boldsymbol{Q}_k(\boldsymbol{v}_{t+2} - \boldsymbol{v}_{t+1}) \\ &= \beta\delta \boldsymbol{Q}_k(\boldsymbol{v}_{t+2} - \boldsymbol{m}_{K,t+1} - \boldsymbol{u}_K - \delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\boldsymbol{v}_{t+2}) \\ &= \beta\delta \boldsymbol{Q}_k(I - \delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1})\boldsymbol{v}_{t+2} - \beta\delta \boldsymbol{Q}_k(\boldsymbol{m}_{K,t+1} + \boldsymbol{u}_K), \end{aligned} \tag{29}$$

where $t \leq T-1$ and, if $t = T-1$, $\boldsymbol{v}_{t+2} = \boldsymbol{v}_{T+1} = 0$ by definition. The first equation holds by definition, the second equality holds by the stationary of the flow utility, and the third equality is obtained by plugging in the definition of the *ex ante* value function $\boldsymbol{v}_{t+1}$ from

Equation (8).[15] I is a $J \times J$ identity matrix. Note that both the transition matrix $\boldsymbol{Q}_k$ and the aggregated transition matrix $\bar{\boldsymbol{Q}}_{t+1}$ can be identified from the data directly. The adjustment term $\boldsymbol{m}_{K,t+1}$ is a known function of the distribution of the error terms ϵ. The remaining unknowns are the two discount factors δ and β, the flow utility $\boldsymbol{u}_K$, and the future *long-run* value function $\boldsymbol{v}_{t+2}$, which is again a function of the discount factors and the flow utility.

Next, we rewrite the recursive relationship between the *perceived long-run value function*

$$\boldsymbol{v}_t = \begin{cases} \boldsymbol{m}_{K,t} + \boldsymbol{u}_K + \delta(1-\beta)\bar{\boldsymbol{Q}}_t \boldsymbol{v}_{t+1}, & \text{for } t < T \\ \boldsymbol{m}_{K,T} + \boldsymbol{u}_K & \text{for } t = T \end{cases}, \tag{30}$$

which depends on the flow utility of the reference action and the two discount factors. This is because the choice-specific value function $w_K(x_t)$ is the same as $u_K(x_t)$ since K is a terminating action. If one observes the agent's behavior up to the final period T, one can simulate the perceived long-run value function by collecting the flow utility in every period up to the final period and thus fully represent the perceived long-run value function as a function of the observed CCPs, the discount factors, and the flow utility $\boldsymbol{u}_K$. Furthermore, if one imposes the normalization assumption that $u_{K,t} = 0$, one can express the perceived long-run value function as a function of the discounting parameters, which allows identification of both β and δ directly.

Instead of assuming that the data is available up to the final period and requiring the normalization of the flow utility, we exploit the recursive structure of the perceived long-run value function and concentrate all variation of the CCPs into functions of β and δ and the utility associated with the terminating action. As before, our goal is to identify both discount factors and the flow utility jointly.

To do this, we first express the perceived long-run value function $\boldsymbol{v}_{t+2}$ as a function of the reference flow utility function $\boldsymbol{u}_K$ and the discount factors, based on Equation (29). That is, if the rank condition from Assumption 5 holds, we have

$$\boldsymbol{v}_{t+2} \equiv h_{t+1,t}(\beta, \delta, \boldsymbol{u}_K) = \left(I - \delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)^{-1} \left(\frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{t+1}) + \boldsymbol{m}_{K,t+1} + \boldsymbol{u}_K\right). \tag{31}$$

Note that $Q_k(I - \delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1})$ has full rank, because $(I - \delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1})$ is full rank if $\delta(1-\beta) \neq 1$, which is the case as long as there is some form of discounting. Consequently, we can uniquely recover the perceived long-run value function if both the exponential discounting and the present-bias parameters and the utility of the terminating action are known.

We can then exploit the recursive relationship in the perceived long-run value function

[15] $\boldsymbol{m}_{K,t+1}$ and $\boldsymbol{u}_K$ stack $m_K(p_{t+1}(x))$ and $u_K(x)$ for all state values x, respectively.

specified in Equation (30) and rewrite it such that it only depends on the discount factors and the flow utility from the terminating action. All previous manipulations use the conditions implied by the model to eliminate the other unknown primitives from this value function. Finally, by combining Equations (30) and (31), we obtain the following key equation:

$$h_{t,t-1}(\beta, \delta, \boldsymbol{u}_K) = \boldsymbol{m}_{K,t+1} + \boldsymbol{u}_K + \delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}h_{t+1,t}(\beta, \delta, \boldsymbol{u}_{\boldsymbol{K}}). \tag{32}$$

Note that this condition requires observing three consecutive periods of CCPs. It includes J equations because there are J potential states of x; and it invovles $1+1+J$ unknowns, i.e., the present-bias parameter β, the exponential discount parameter δ, and the vector of flow utilities from the terminating action $\boldsymbol{u}_{\boldsymbol{K}}$. Consequently, we are not able to identify all three components from Equation (32) when we use only CCPs from three periods.

In order to proceed, assume for now that the discount factors β and δ are known. In this case we can identify the flow utility associated with the terminating action when the equilibrium CCPs are non-stationary. Therefore, we impose the following assumption.

Assumption 8 *There exists an action k such that we have $p_{k,t}(x) \neq p_{k,t+1}(x)$, for all x.*

Intuitively, this assumption requires that there is enough variation of the CCPs for at least one action over time. This assumption is a higher-level assumption in the sense that it is imposed on endogenous components of the model instead of the model primitives. However, this assumption is easy to satisfy by construction due to the finite horizon framework, which usually assumes that the continuation value in the final period is zero. Moreover, it is directly testable from the data.

Once the flow utility associated with the terminating action is identified, the identification of the flow utilities associated with any other action j can be achieved analogously by using the respective value contrasts between u_j and the reference utility u_K. We summarize our identification results for the flow utility functions in the following proposition.

Proposition 3 *If*

1. *the discount factors β and δ are known, and*
2. *Assumptions 1 to 5 and 8 hold,*
3. *all agents are sophisticated,*

then all flow utility functions are identified using any three consecutive periods of data.

We relegate our proof of Proposition 3 to Appendix A. The arguments in the proof are closely related to those from the literature on using the presence of a terminating action to identify the flow utilities in an exponential discounting framework under the assumption that the discount factor is known, see, for example, Bajari et al. (2016) and Blevins et al. (2020).

Proposition 3 implies that any three consecutive periods of data can identify the flow utility u_K as a closed-form function of the two discount factors β and δ and the observed CCPs. That is, we can explicitly represent the flow utility as $u_K = \Upsilon(\beta, \delta, \boldsymbol{p}_t, \boldsymbol{p}_{t+1}, \boldsymbol{p}_{t+2})$, where $\Upsilon(\cdot)$ has a closed-form expression. Consequently, if we observe more than three periods of data, Proposition 3 provides overidentifying restrictions for the two discount factors. Specifically, if we observe data for four consecutive periods, i.e., we can compute $\{\boldsymbol{p}_t, \boldsymbol{p}_{t+1}, \boldsymbol{p}_{t+2}, \boldsymbol{p}_{t+3}\}$ directly from the data, we can identify and estimate two sets of flow utilities separately by using CCPs from any three consecutive, i.e., $u_K^1 = \Upsilon(\beta, \delta, \boldsymbol{p}_t, \boldsymbol{p}_{t+1}, \boldsymbol{p}_{t+2})$ and $u_K^2 = \Upsilon(\beta, \delta, \boldsymbol{p}_{t+1}, \boldsymbol{p}_{t+2}, \boldsymbol{p}_{t+3})$. The equality of u_K^1 and u_K^2, which comes from the stationarity assumption on the flow utility, is essential for constructing the identifying restrictions on the two discounting parameters. Specifically, we can write the restrictions as

$$R^s(\beta, \delta, \boldsymbol{p}_t, \boldsymbol{p}_{t+1}, \boldsymbol{p}_{t+2}, \boldsymbol{p}_{t+3}) \equiv \Upsilon(\beta, \delta, \boldsymbol{p}_t, \boldsymbol{p}_{t+1}, \boldsymbol{p}_{t+2}) - \Upsilon(\beta, \delta, \boldsymbol{p}_{t+1}, \boldsymbol{p}_{t+2}, \boldsymbol{p}_{t+3}), \tag{33}$$

which equals zero at the true values of the two discounting parameters. Note that this condition is not linear in the two discount factors and has J restrictions but only two unknowns. Consequently, if the gradient of the J restrictions, which is of size $J \times 2$, has a rank of 2 at the true parameter values, we can locally identify the two discounting parameters. We formalize this rank condition in the following assumption.

Assumption 9 *The gradient of the J restrictions, denoted as $\nabla R^s(\beta, \gamma) \equiv \begin{bmatrix} \frac{\partial R}{\partial \beta} \\ \frac{\partial R}{\partial \gamma} \end{bmatrix}$, has a rank of 2 at the true parameter values.*

Local identification of the model primitives then follows, and we summarize this result in the following Proposition.

Proposition 4 *If Assumptions 1 to 5, 8, and 9 are satisfied, then all flow utility functions, the discount factor δ, and the present-bias parameter β are (locally) identified using any four consecutive periods of data.*

It is worth noting that Proposition 4 does not rely on the zero continuation value assumption in the final period. Once the flow utility, the discount factor, and the present-bias parameter are identified, one can identify the continuation value, which requires the conventional normalization assumption. Note that when we observe the agents' actions in the intermediate periods instead of the final three periods, we cannot globally identify the model without a normalization condition. However, if we are willing to impose an arguably milder normalization assumption, namely, that the flow utility associated with the terminating action is known for two values of the state x, we can fully identify both the product of $\beta\delta$ and the flow utility $\boldsymbol{u}_K$ for other states. This assumption is milder than the ones imposed in the existing literature. For instance, Mahajan, Michel, and Tarozzi (2020) assume that the whole vector of $\boldsymbol{u_K}$ is known; Abbring, O. Daljord, and Iskhakov (2019) assume that $\boldsymbol{u_K} = 0$.

Identification problems in the naive agent case. In the following, we illustrate the challenges for identification in the naive agent framework, when the final periods are not observed. As a starting point, consider the stacked log odds ratios

$$\boldsymbol{\phi}_{kK}(\boldsymbol{p}_t) = \boldsymbol{w}_{k,t} - \boldsymbol{w}_{K,t} = \boldsymbol{u}_{k,t} - \boldsymbol{u}_{K,t} + \beta\delta\boldsymbol{Q}_k\boldsymbol{v}_{t+1}(\boldsymbol{p}^t_{t+1}). \tag{34}$$

The key difference between these log odds ratios and the ones described in Equation (28) is that the perceived long-run value function is computed based on the future self's optimal behavior $\boldsymbol{p}^t_\tau$ in the perception of the current self, where $\tau \geq t+1$. In contrast to the sophisticated agent, the naive agent believes that the future self is time-consistent. Therefore, we can write the perceived long-run value function as

$$\boldsymbol{v}_{t+1}(\boldsymbol{p}^t_{t+1}) \;\; = \;\; \boldsymbol{m}_K(\boldsymbol{p}^t_{t+1}) + \boldsymbol{z}_{K,t+1} = \boldsymbol{m}_K(\boldsymbol{p}^t_{t+1}) + \boldsymbol{u}_K. \tag{35}$$

The second equality holds because action K is a terminating action so that the continuation value is zero. As in the previous subsection, the presence of a terminating action simplifies the representation of the perceived long-run value function such that it can be fully characterized by the flow utility associated with the terminating action $\boldsymbol{u}_K$ and the adjustment term associated with the probability of choosing this action $\boldsymbol{m}_K(\boldsymbol{p}^t_{t+1})$.

The key difficulty for identification of this model is that we cannot recover $\boldsymbol{p}^t_{t+1}$ from the data as in the sophisticated agent framework. Consequently, the identification strategy developed for the sophisticated agent framework is not readily applicable. Because of this difficulty, the existing literature almost exclusively focuses on models with sophisticated agents. To the best of our knowledge the only exception is Mahajan, Michel, and Tarozzi (2020), who study the identification of naive agents' discount factors using purposefully collected data on the final three periods.

4 Estimation

In this section, we propose two estimators for the discount factors and the flow utilities. First, the *sequential estimator* follows closely our identification strategy discussed in the previous section. It estimates the discount factors in a first step and the flow utilities in a second step. The *joint estimator* estimates the discount factors and the parametrized flow utilities in one step by maximizing likelihood.

Throughout, we assume that we observe data on actions and states $\{a_{it}, x_{it}\}_{i=1,\dots,N,t=t_1,t_1+1,\dots,t_1+l}$, where t_1 is the first observed period and $t_1 + l$ denotes the last observed period, which may equal T or may be strictly less than T, i.e., the action in the final period is not observed. Moreover, we impose that $l >= 2$ when $t_1 + l = T$, or $l >= 3$ when $t_1 + l < T$, to make sure that the identification assumptions are satisfied.

In all cases, we first estimate the equilibrium CCPs for each action a and observed state x via a simple frequency estimator.[16] That is,

$$\hat{p}_t(a_t = 1|x_t = x) = \frac{\sum_i I(a_{it} = 1, x_{it} = x)}{\sum_i I(x_{it} = x)}, \quad t = t_1, t_1 + 1, \dots, t_1 + l. \tag{36}$$

The transition matrix of the observed state x is estimated analogously. Note that both the equilibrium CCPs and the transition matrix are estimated consistently when the sample size goes to infinity.

4.1 Sequential Estimation

For the case of the sophisticated agent with the final periods observed, we can estimate the two discount factors separately via an OLS estimator. For the case in which the final periods are not observed, or the agents are naive, we present a minimum distance (MD) estimator. Once the two discount factors are estimated, the flow utilities can be estimated non-parametrically in a second step.

If the last three periods are observed, we can estimate the two discount factors for the sophisticated agent via the following OLS estimator, which is based on the identification condition in Equation (22).

$$\begin{bmatrix} \hat{\delta}\hat{\beta} \\ \hat{\delta} \end{bmatrix} = \left(\Omega'(\hat{\boldsymbol{p}}_{T-2}, \hat{\boldsymbol{p}}_{T-1}, \hat{\boldsymbol{p}}_T)\Omega(\hat{\boldsymbol{p}}_{T-2}, \hat{\boldsymbol{p}}_{T-1}, \hat{\boldsymbol{p}}_T)\right)^{-1} \Omega'(\hat{\boldsymbol{p}}_{T-2}, \hat{\boldsymbol{p}}_{T-1}, \hat{\boldsymbol{p}}_T)\Delta\boldsymbol{\phi}_{kK}(\hat{\boldsymbol{p}}_{T-1}). \tag{37}$$

If we observe the final three periods, the two discount factors for naive agents can be

[16] In principle, one can employ any other consistent first-step estimator.

estimated using a minimum distance estimator following Equation (27).[17] That is,

$$
\begin{aligned}
\{\hat{\beta}, \hat{\delta}\} &= \arg min_{\beta,\delta} || \frac{1}{\beta\delta} \boldsymbol{Q}_k^{-1} \left(\boldsymbol{\phi}_{kK}(\hat{\boldsymbol{p}}_{T-2}) - \boldsymbol{\phi}_{kK}(\hat{\boldsymbol{p}}_T) \right) \\
&- \left[log \left(\sum_k exp \left(\boldsymbol{\phi}_{kK}(\hat{\boldsymbol{p}}_T) - \frac{1}{\beta} \Delta \boldsymbol{\phi}_{kK}(\hat{\boldsymbol{p}}_T) \right) \right) + log(\hat{\boldsymbol{p}}_{KT}) - \frac{1}{\beta\delta} \boldsymbol{Q}_k^{-1} \Delta \boldsymbol{\phi}_{kK}(\hat{\boldsymbol{p}}_T) \right] ||_2,
\end{aligned}
$$

where $||\cdot||$ denote the $L2$-norm.

If the final periods are not observed in the data, the two discount factors of the sophisticated agent can be estimated using a similar MD estimator, based on the identification restrictions in Equation (33). That is,

$$
\{\hat{\beta}, \hat{\delta}\} = \arg min_{\beta,\delta} ||R^s(\beta, \delta, \hat{\boldsymbol{p}}_t, \hat{\boldsymbol{p}}_{t+1}, \hat{\boldsymbol{p}}_{t+2}, \hat{\boldsymbol{p}}_{t+3})||_2.
$$

The asymptotic properties of such two-step estimators are well established, see, for example, Pesendorfer and Schmidt-Dengler (2008). Under standard regularity conditions, the two-step estimators proposed above are consistent and asymptotically normally distributed when the sample size goes to infinity.

4.2 Joint Maximum Likelihood Estimation

In practice, estimating both the discount factors and the flow utilities sequentially and non-parametrically is often too demanding on the data. Especially with limited sample sizes, researchers often specify and estimate the payoff function parametrically. In this subsection, we present a maximum likelihood estimator that recovers the discount factors and the parameterized flow utility jointly.

Let θ denote all parameters, i.e., $\theta \equiv \{\beta, \delta, \theta_j\}$, where θ_j captures the parameters in the flow utility associated with action j. The maximum likelihood estimator is standard and given by

$$
\hat{\theta} = \arg max_\theta \sum_i \sum_t log \left(p_t(a_{it}|x_{it}; \theta) \right). \tag{38}
$$

The asymptotic properties of this estimator are established in Theorems 9 and 11 of John (1988). That is, as the number of individuals becomes large, the sequence of maximum likelihood estimators are consistent and asymptotically normally distributed.

[17] This estimator is similar in spirit to the one proposed by Abbring, O. Daljord, and Iskhakov (2019) for the sophisticated agent.

5 Monte Carlo Simulations

In this section, we analyze a stylized model setup to illustrate that our proposed estimators work well in simulations.[18] First, we introduce the model setup and its parametrization.[19] Afterwards, we discuss our results from applying the estimators discussed in Section 4 to simulated data. We first illustrate the finite sample properties of the discount factor estimators discussed in Section 4.1. Not surprisingly, these estimators requires a large sample size in order to estimate the parameters precisely. Therefore, we document the performance of the full model estimation, i.e., discount factors and flow utilities jointly, as discussed in Section 4.2. Throughout, the maximum likelihood estimators perform very well even with modest sample sizes. Lastly, we illustrate the importance of avoiding the normalization of the flow utility for both estimation and counterfactuals. In order to keep the illustration concise, we focus on the settings in which the last three periods are observed in the data.

5.1 Model Setup

We consider a stylized solar panel adoption problem similar to the application in De Groote and Verboven (2019). In each period t, agent i can choose to adopt a solar panel, i.e., $a_{it} = 1$, or wait for next period $t + 1$, i.e., $a_{it} = 0$. Once the agent makes the adoption decision, she is out of the market and never considers the adoption decision again. The agent observes the price-adjusted quality of the solar panel, which is the only state variable $x_t \in Supp(\mathcal{X}) = \{2, 3, 7, 9\}$. We set the true discount factor to $\delta = 0.8$ and the present-bias parameter to $\beta = 0.4$. We assume that naive agents have the same discount factors as the sophisticated agents, but naive agents believe that they will be time-consistent in the future. We specify the flow utility as

$$u(x, a) = \begin{cases} \theta_1 + \theta_2 x & \text{if } a_i = 1 \\ \theta_3 + \theta_4 x & \text{if } a_i = 0, \end{cases} \tag{39}$$

with $\theta_1 = 2.5, \theta_2 = 0.7, \theta_3 = 0, \theta_4 = 1$. We set the transition matrix of the state variable to

$$Q_0(x'|x) = \begin{bmatrix} 0.4800 & 0.2400 & 0.1600 & 0.1200 \\ 0.2143 & 0.4286 & 0.2143 & 0.1429 \\ 0.1429 & 0.2143 & 0.4286 & 0.2143 \\ 0.1200 & 0.1600 & 0.2400 & 0.4800 \end{bmatrix}. \tag{40}$$

[18] In Appendix B.1 we provide applied researchers with additional step-by-step guidance on how to implement our identification strategy in practice.

[19] In Appendix B.1, we analyze and visualize the true decision strategies for all three types of agents (time-consistent, sophisticated, and naive).

Given the model primitives, we can solve for the equilibrium CCPs. Figure 3 in Appendix B.1 presents the adoption rates for different types of agents (time-consistent, sophisticated, and naive) in different states.[20] As x becomes larger, the adoption rate decreases since θ_2 is smaller than θ_4, so that the utility from waiting becomes relatively larger. For the same x, the adoption rate becomes larger as time advances. This is because the adoption is a terminating action and there is no future payoff; therefore, agents may choose to wait until the final period to realize a higher lifetime utility even though the flow utility from adopting is higher. Furthermore, the adoption rate is higher for sophisticated and naive agents and the adoption rate for sophisticated agents is the highest. This is consistent with present-biased behavior: Since it is better for time-consistent agent to wait until the final period, present-biased agents prefer today's payoff more and adopt earlier.

5.2 Simulation Results

We simulate the data based on the DGP presented in Section 5.1 and use 100 simulation runs for each scenario. Table 6 in Appendix B.1 summarizes the estimation results from our OLS estimator for the case of the sophisticated agent and sample sizes of 50,000, 100,000, and 1,000,000. Not surprisingly, the performance of the estimators improves with larger sample sizes, i.e., the estimates of the discount factors are closer to the true values, and the standard error shrinks. It is worth noticing that the OLS estimator mainly exploits the variation in CCPs over time, while aggregating over all individuals. Such an estimator is simple but requires a lot of over-time variation in the data. Similar results using similar sample sizes are reported by Abbring and Ø. Daljord (2020a).

For the naive agent case, we simulate the data using the true CCPs for sample sizes of 50,000, 100,000, and 1,000,000. We estimate the two discount factors using our MD estimator. Table 7 in Appendix B.1 presents the mean and standard deviation from these simulations. Both discount factors are estimated precisely for all of our sample sizes.

When estimating discount factors and flow utilities jointly using maximum likelihood, we use sample sizes of 5,000, 10,000, and 20,000 for both the sophisticated and the naive agent settings. Tables 8 and 9 in Appendix B.1 present the results for the sophisticated agent and the naive agent, respectively. In both settings, the MLE estimator performs well even with a moderate sample size of 5,000.

To understand the importance of the normalization, which is typically imposed in the existing literature, we also estimate the full model under a normalization condition on the flow utility. In this case, the maximum likelihood estimator is modified to incorporate the

[20]For illustration purposes, we present the adoption CCPs only for the last 4 periods of the problem.

normalization as the follows:

$$\hat{\theta}^{normalization} = \arg max_{\theta;\theta_1^u=0} \sum_i \sum_t log\left(p_t(a_{it}|x_{it};\theta)\right). \tag{41}$$

The estimation results when normalizing the utility associated with the adoption action are displayed in Table 10 in Appendix B.1.[21] In our maximum likelihood estimation, which estimates utilities and discount factors jointly, the normalization of the flow utility leads to substantially biased estimates of the discount factors.

More importantly, we study the impact of the normalization on counterfactual analyses. Given the model estimates, we simulate the counterfactual outcomes of an adoption subsidy using both models, the one with and the one without the flow utility normalization. In the normalized setting, the utility from adoption is set to zero; therefore, we implement the counterfactual by raising the utility level by 0.5. That is, θ_1 is decreased by 0.3 in the case of normalizing the non-adoption decision, and θ_3 is increased by 0.5 in the case of normalizing the adoption decision.

We plot the counterfactual CCPs with and without normalization in Figure 4 in Appendix B.1.[22] We can see that that, under a subsidy for adopting, the model without normalization increases the adoption rate only slightly compared to the status quo, while the model with normalization raises the adoption rate much more in the first few periods. Overall the predictions from a normalized model differ strikingly from the non-normalized model. This results is in line with the recent literature on conducting counterfactuals in dynamic models, see, for example, Kalouptsidi, Scott, and Souza-Rodrigues (2021).

6 Conclusion

In this paper, we study the identification of DDCMs with hyperbolic discounting. We focus on the economically relevant class of DDCMs with a finite horizon in which agents can choose a terminating action to end the decision problem. Under the assumption of a stationary flow utility we provide novel identification results for both sophisticated and naive agents' discount factors and their flow utilities. Our identification strategy exploits the recursive structure of the DDCM and variation in the CCPs over time. Compared to existing identification strategies our approach has several advantages. First, we do not require to observe the final period to identify the parameters for the sophisticated agent

[21] We also estimate the model normalizing the non-adoption action. The results are qualitatively similar and available upon request.

[22] We plot the simulated CCPs for ten periods to see the long-run effects. The figures are based on a sample size of 5,000.

as long as we observe four consecutive periods of data. Second, we show identification of the naive agent's parameters without any special data requirement, such as data on agents' beliefs. However, for the naive agent we require that the final three periods of data are observed. Lastly, we avoid having to normalize the flow utility of a reference action for both the sophisticated and the naive agent. Recent research discusses that such a normalization often biases counterfactual simulations, see, for example, Kalouptsidi, Scott, and Souza-Rodrigues (2021).

Based on our constructive identification proof, we propose two tractable estimators. If the final three periods of data are observed, the discount parameters of the sophisticated agent can be recovered using a simple OLS estimator. If the final periods are not observed or if the agent is naive, we obtain polynomial moment conditions that form the basis of a minimum distance estimator similarly to the one proposed by Abbring, O. Daljord, and Iskhakov (2019). Both estimators perform well in Monte Carlo simulations. Our simulations also indicate that more restrictive estimation approaches, such as the ones that impose an artificial normalization of the flow utility, generally result in biased counterfactual policy predictions.

Many applications in industrial organization, labor or health, in particular, decisions about long-term financial products, such as mortgages, or technology adoption and investment decisions fit into our framework. Given the rising interest in empirical models with hyperbolic discounting, our identification and estimation strategy provides an important step to empirically investigate a broad range of important dynamic problems in a more flexible way than is possible with existing approaches.

References

Abbring, Jaap H. and Øystein Daljord (2020a). "A comment on "Estimating dynamic discrete choice models with hyperbolic discounting" by Hanming Fang and Yang Wang". *International Economic Review* 61.2, pp. 565–571.

— (2020b). "Identifying the discount factor in dynamic discrete choice models". *Quantitative Economics* 11.2, pp. 471–501.

Abbring, Jaap H., Oystein Daljord, and Fedor Iskhakov (2019). "Identifying Present-Biased Discount Functions in Dynamic Discrete Choice Models". *Working Paper.*

Aguirregabiria, Victor and Junichi Suzuki (2014). "Identification and counterfactuals in dynamic models of market entry and exit". *Quantitative Marketing and Economics* 12.3, pp. 267–304.

An, Yonghong, Yingyao Hu, and Ruli Xiao (2021). “Dynamic decisions under subjective expectations: A structural analysis”. *Journal of Econometrics* 222.1, pp. 645–675.

Arcidiacono, Peter and Robert A. Miller (2011). “Conditional choice probability estimation of dynamic discrete choice models With unobserved heterogeneity”. *Econometrica* 79.6, pp. 1823–1867.

Bajari, Patrick et al. (2016). “Identification and semiparametric estimation of a finite horizon dynamic discrete choice model with a terminating action”. *Quantitative Marketing and Economics* 14.4, pp. 271–323.

Blevins, Jason R. et al. (2020). “A dynamic discrete choice model of reverse mortgage borrower behavior”. *International Economic Review* 61.4, pp. 1437–1477.

Chan, Marc K. (2017). “Welfare dependence and self-control: An empirical analysis”. *Review of Economic Studies* 84.4, pp. 1379–1423.

Colas, Mark, Sebastian Findeisen, and Dominik Sachs (2021). “Optimal need-based financial aid”. *Journal of Political Economy* 129.2, pp. 492–533.

Daljord, Øystein, Denis Nekipelov, and Minjung Park (2019). “Comments on “Identification and semiparametric estimation of a finite horizon dynamic discrete choice model with a terminating action””. *Quantitative Marketing and Economics* 17.4, pp. 439–449.

De Groote, Olivier and Frank Verboven (2019). “Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems”. *American Economic Review* 109.6, pp. 2137–2172.

Eckstein, Zvi and Kenneth I. Wolpin (1999). “Why youths drop out of high school: The impact of preferences, opportunities, and abilities”. *Econometrica* 67.6, pp. 1295–1339.

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, pp. 565–596.

Heidhues, Paul and Philipp Strack (2021). “Identifying present bias from the timing of choices”. *American Economic Review* 111.8, pp. 2594–2622.

Hotz, V. Joseph and Robert A. Miller (1993). “Conditional choice probabilities and the estimation of dynamic models”. *The Review of Economic Studies* 60.3, pp. 497–529.

Hu, Yingyao (2008). “Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution”. *Journal of Econometrics* 144.1, pp. 27–61.

John, Rust (1988). “Maximum likelihood estimation of discrete control processes”. *SIAM journal on control and optimization* 26.5, pp. 1006–1024.

Kalouptsidi, Myrto (2014). “Time to build and fluctuations in bulk shipping”. *American Economic Review* 104.2, pp. 564–608.

Kalouptsidi, Myrto, Paul T. Scott, and Eduardo Souza-Rodrigues (2021). "Identification of counterfactuals in dynamic discrete choice models". *Quantitative Economics* 12.2, pp. 351–403.

Laibson, David (1997). "Golden eggs and hyperbolic discounting". *Quarterly Journal of Economics* 112.2, pp. 442–477.

Levy, Matthew and Pasquale Schiraldi (2023). "Identification of intertemporal preferences from choice set variation". *SSRN Working Paper*.

Magnac, Thierry and David Thesmar (2002). "Identifying dynamic discrete decision processes". *Econometrica* 70.2, pp. 801–816.

Mahajan, Aprajit, Christian Michel, and Alessandro Tarozzi (2020). *Identification of time-inconsistent models: The case of insecticide treated nets*. Working Paper 27198. National Bureau of Economic Research.

Manski, Charles F, Daniel McFadden, et al. (1981). *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press Cambridge, MA.

Norets, A. and X. Tang (2014). "Semiparametric inference in dynamic binary choice models". *Review of Economic Studies* 81.3, pp. 1229–1262.

O'Donoghue, Ted and Matthew Rabin (1999). "Doing it now or later". *American Economic Review* 89.1, pp. 103–124.

Pesendorfer, Martin and Philipp Schmidt-Dengler (2008). "Asymptotic least squares estimators for dynamic games". *The Review of Economic Studies* 75.3, pp. 901–928.

Rothenberg, Thomas J (1971). "Identification in parametric models". *Econometrica: Journal of the Econometric Society*, pp. 577–591.

Rust, John (1994). "Structural estimation of Markov decision processes". *Handbook of Econometrics* 4, pp. 3081–3143.

Tsubota, Taiga (2021). "Identifying Dynamic Discrete Choice Models with Hyperbolic Discounting". *arXiv preprint arXiv:2111.10721*.

Appendices

The Appendix consists of two parts: the detailed proofs for the lemmas and theorems presented in the main text and additional details on our Monte Carlo simulations, that can also serve as guidance for empirical researchers to better understand the identification assumptions and their implications for applied work.

A Proofs

This section provides all proofs relegated from the main text.

Proof of Proposition 3. From Equation (32) we have the following conditions that involve the two discount factors and the utility of the terminating action:

$$
\begin{aligned}
& \left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_t\right)^{-1}\left(\frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_t)+\boldsymbol{m}_{K,t}+\boldsymbol{u}_K\right) \\
= \;& \boldsymbol{m}_{K,t+1}+\boldsymbol{u}_K+\delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)^{-1}\left(\frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{t+1})+\boldsymbol{m}_{K,t+1}+\boldsymbol{u}_K\right) \\
= \;& \boldsymbol{m}_{K,t+1}+\boldsymbol{u}_K-\left[I-\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)^{-1}\right]\left(\frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{t+1})+\boldsymbol{m}_{K,t+1}+\boldsymbol{u}_K\right) \\
\leftrightarrow \;& \left[\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_t\right)^{-1}-\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)^{-1}\right]\boldsymbol{u}_K \\
= \;& \boldsymbol{m}_{K,t+1}-\left[I-\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)^{-1}\right]\left(\frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{t+1})+\boldsymbol{m}_{K,t+1}\right) \\
- \;& \left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_t\right)^{-1}\left(\frac{1}{\beta\delta}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_t)+\boldsymbol{m}_{K,t}\right) \\
\equiv \;& H(\beta,\delta,\boldsymbol{Q}_k,\boldsymbol{p}_{t-1},\boldsymbol{p}_t,\boldsymbol{p}_{t+1}), \quad k\neq K,
\end{aligned} \tag{42}
$$

where the component from the right-hand side can be directly computed if we know β and δ. We can further simply the coefficient in front of the flow utility as follows:

$$
\begin{aligned}
& \left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_t\right)^{-1}-\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)^{-1} \\
= \;& \left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_t\right)^{-1}\left[\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)-\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_t\right)\right]\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)^{-1} \\
= \;& \delta(1-\beta)\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_t\right)^{-1}\left[\bar{\boldsymbol{Q}}_t-\bar{\boldsymbol{Q}}_{t+1}\right]\left(I-\delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)^{-1}.
\end{aligned}
$$

Therefore, as long as $\left[\bar{\boldsymbol{Q}}_t - \bar{\boldsymbol{Q}}_{t+1}\right]$ is full rank, we can uniquely solve for the flow utility associated with the terminating action in closed-form:

$$\begin{aligned}
\boldsymbol{u}_K &= \frac{1}{\delta(1-\beta)}\left(I - \delta(1-\beta)\bar{\boldsymbol{Q}}_{t+1}\right)\left[\bar{\boldsymbol{Q}}_t - \bar{\boldsymbol{Q}}_{t+1}\right]^{-1}\left(I - \delta(1-\beta)\bar{\boldsymbol{Q}}_t\right) H(\beta,\delta,\boldsymbol{Q}_k,\boldsymbol{p}_{t-1},\boldsymbol{p}_t,\boldsymbol{p}_{t+1}) \\
&\equiv \Upsilon(\beta,\delta,\boldsymbol{p}_{t-1},\boldsymbol{p}_t,\boldsymbol{p}_{t+1}) \qquad (43)
\end{aligned}$$

We can show that $\left[\bar{\boldsymbol{Q}}_t - \bar{\boldsymbol{Q}}_{t+1}\right]$ is indeed of full rank given Assumptions 4 and 5. Recall that $\bar{\boldsymbol{Q}}_t$ is a $J \times J$ matrix with element (i,j) equal to $\sum_k Q_k(x_{t+1} = j | x_t = i) p_{k,t}(x_t = i)$. Consequently, the difference in the compound state evolution can be written as

$$\begin{aligned}
\bar{\boldsymbol{Q}}_t - \bar{\boldsymbol{Q}}_{t+1} &= \sum_k diag(\boldsymbol{p}_{k,t})\boldsymbol{Q}_k - \sum_k diag(\boldsymbol{p}_{k,t+1})\boldsymbol{Q}_k \\
&= \sum_k \left[diag(\boldsymbol{p}_{k,t}) - diag(\boldsymbol{p}_{k,t+1})\right]\boldsymbol{Q}_k \\
&= \begin{bmatrix} diag(\boldsymbol{p}_{1,t}) - diag(\boldsymbol{p}_{1,t+1}) & \cdots & diag(\boldsymbol{p}_{K,t}) - diag(\boldsymbol{p}_{k,t+1}) \end{bmatrix} \times \begin{bmatrix} \boldsymbol{Q}_1 \\ \vdots \\ \boldsymbol{Q}_K \end{bmatrix} \\
&\equiv D \times E,
\end{aligned}$$

where $diag(\boldsymbol{p}_{k,t})$ is a $J \times J$ diagonal matrix with the (j,j)-th element being $p_{k,t}(x = j)$, D is a matrix of sizes $J \times JK$, and E is a matrix of sizes $JK \times J$. $rank(E) = J$ because of Assumption 5, and $rank(D) = J$ because of Assumption 5. Therefore, $\bar{\boldsymbol{Q}}_t - \bar{\boldsymbol{Q}}_{t+1}$ is of full rank by Sylvester's rank inequality. That is,

$$rank(D \times E) \geq rank(D) + rank(E) - J = J. \qquad (44)$$

Once the flow utility associated with the terminating action is identified, we can identify other flow utilities from Equation (28).

Proof of Proposition 4. From Proposition 3, we can identify the flow utility associated with the terminating action in closed-form given that the hyperbolic discounting parameters can be identified using three periods of data. If we have four periods of data, this provides over-identification restrictions on this flow utility, which we can exploit to identify the two unknown parameters. Specifically, with four periods of data, we can obtain the equilibrium CCPs $\boldsymbol{p}_t, \boldsymbol{p}_{t+1}, \boldsymbol{p}_{t+2}, \boldsymbol{p}_{t+3}$, which provides the following over-identification restrictions:

$$\Upsilon(\beta,\delta,\boldsymbol{p}_t,\boldsymbol{p}_{t+1},\boldsymbol{p}_{t+2}) = \Upsilon(\beta,\delta,\boldsymbol{p}_{t+1},\boldsymbol{p}_{t+2},\boldsymbol{p}_{t+3}).$$

B Guidance for Empirical Work

In this subsection we provide a road map for applied researchers on how to take our identification strategy to real-world data using our Monte Carlo setup from Section 5 as an illustrative example. To simplify the illustration and to avoid finite sample error from the first step, we assume that we know the population CCPs and the state transition matrix.[23] We discuss sophisticated and naive agents separately. Given that all our technical assumptions are rank conditions, we rely on the singular values of the matrix to verify that all regularity conditions imposed by our assumptions are satisfied.[24]

Sophisticated agent. Our first step is to verify the full rank condition on the state transition matrix Q_0 imposed by Assumption 5. Table 1 displays the singular values for matrix $\boldsymbol{Q}_0$. We see that the number of positive singular values is the same as the number of states for all columns. Therefore, the rank condition in Assumption 5 is satisfied in our example.

SV1	SV2	SV3	SV4
1.0010	0.3923	0.2426	0.1821

Table 1: Singular values for $\boldsymbol{Q}_0$

Second, we verify the rank condition specified in Assumption 6, which is required for identification when we observe the data up to the final period, see Proposition 1. We need to check the rank of matrix $\Omega(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)$. The form of $\Omega(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)$ and its singular values are shown in Table 2. Since two singular values are positive, $rk(\Omega(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)) = 2$ holds.

	col1	col2
$\Omega_{1,\cdot}$	-0.4544	1.4398
$\Omega_{2,\cdot}$	-0.4774	1.5965
$\Omega_{3,\cdot}$	-0.4747	1.8886
$\Omega_{4,\cdot}$	-0.4470	2.0580
SV(Ω)	3.6423	0.1310

Table 2: Ω and its singular values

Once the rank condition is satisfied, we can directly compute the two discount factors

[23]The best way to estimate CCPs and transition matrices in the first step depends on the specific application and the available data.

[24]The singular values of a matrix are the absolute values of its eigenvalues. The rank of a matrix is determined by the number of its positive singular values.

using Equation (22)

$$\begin{bmatrix}\delta\beta \\ \delta\end{bmatrix} = \left(\Omega'(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)\Omega(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)\right)^{-1} \Omega'(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-1}). \quad (45)$$

The computation result is shown in Table 3. The computed discount factors are identical to the true parameters.

$\delta\beta$	δ
0.3200	0.8000

Table 3: Verified discount factors

Third, we examine the condition imposed by Assumption 8. Table 4 presents the differences between the CCP of waiting (action 0) in different time periods.[25] All values are nonzero, which means that the non-zero condition in Assumption 8 holds in our example and both matrices seem well-posed.

	$p_{0,T-3} - p_{0,T-2}$	$p_{0,T-2} - p_{0,T-1}$	$p_{0,T-1} - p_{0,T}$
x=2	0.1372	0.2297	0.3815
x=3	0.0961	0.2130	0.4543
x=7	0.0209	0.0827	0.4837
x=9	0.0083	0.0413	0.3963

Table 4: $p_{k,t} - p_{k,t+1}$

Given that all regularity conditions are satisfied, we can follow Proposition 2 to non-parametrically identify and compute the flow utility u_1 based on Equation (32), using any three consecutive periods of data on the CCPs.[26] We present the associated results in Table 5. Specifically, we compute the payoff u_1 using the true CCPs from period 1 to 3 and from period 2 to 4, separately, see column 3 and 4, respectively. All the computed reference utility utilities are the same as the true ones.

Lastly, we verify that there is a unique solution for the two discount factors from the over-identifying restrictions, as presented in Proposition 3. Note that u_1 can be expressed as a closed-form function of β and δ given the CCPs from any three consecutive periods. Given that we have four periods of data, we can recover u_1 separately using the CCPs in period 1 to 3 and in periods 2 to 4. To verify that there is unique solution for the two discount factors, we check the rank condition specified in Assumption 9.

[25]Since there are only two actions in this example, the CCP differences for other action are the opposite numbers of the values in the table.

[26]Recall that at this stage we assume that the two discount factors are known.

	True u_1	u_1^{123}	u_1^{234}
x=2	3.9000	3.9000	3.9000
x=3	4.6000	4.6000	4.6000
x=7	7.4000	7.4000	7.4000
x=9	8.8000	8.8000	8.8000

Table 5: true u_1 and computed u_1 from three consecutive periods of data

As an alternative check for identification, we can represent the distance between the two sets of utility as a function of the two discount factors:

$$\Delta(\beta, \delta) \equiv |\Upsilon(\beta, \delta, \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) - \Upsilon(\beta, \delta, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4)|, \tag{46}$$

where $|\cdot|$ represents the Euclidean norm. Note that identification means that only the set of the true values of our parameters make this distance is zero. We investigate this graphically, in Figure 1. The "colder points" in this figure, the lower the distance criterion is. The distance criterion is zero only at the true β and δ values.[27]

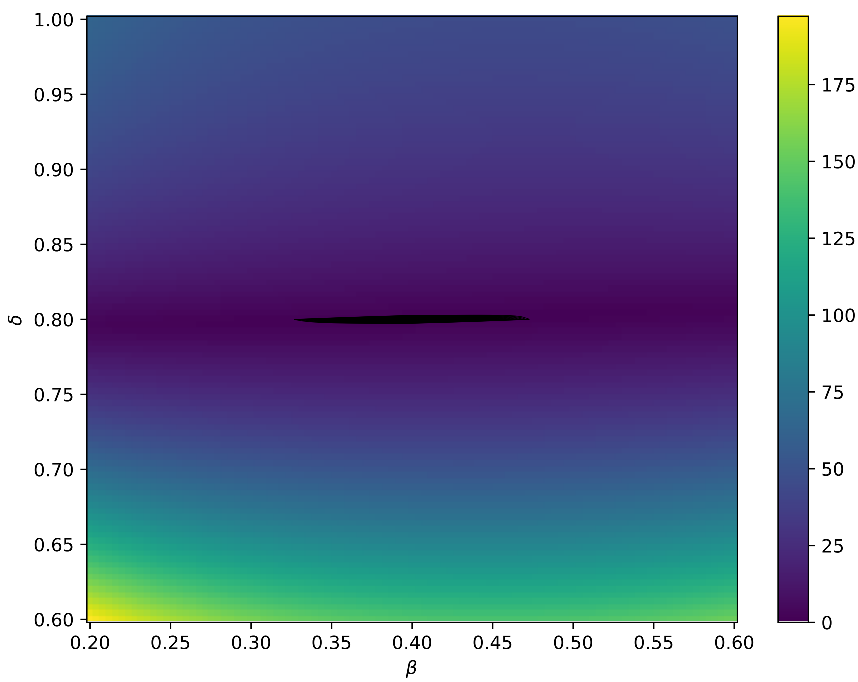

Figure 1: Distance criterion for sophisticated agent using four consecutive periods of data

[27]In Figure 1, the distance $\Delta(\beta, \delta)$ gets smaller along δ fast, but relatively slowly in the β-dimension. Although the figure only shows the values within the true value ± 0.2, the trend holds for the whole range between 0 and 1.

Naive agent. For the naive agent, the identification boils down to whether there is a unique solution to the condition specified in Equation (27). Therefore, we need to check the rank condition specified in Assumption 7, which can be done analogously to the sophisticated agent case. We can also check identification for the naive agent framework using the distance criterion as a function of the discount factors. As for the sophisticated agent case, the distance is minimized at the true values, see Figure 2.

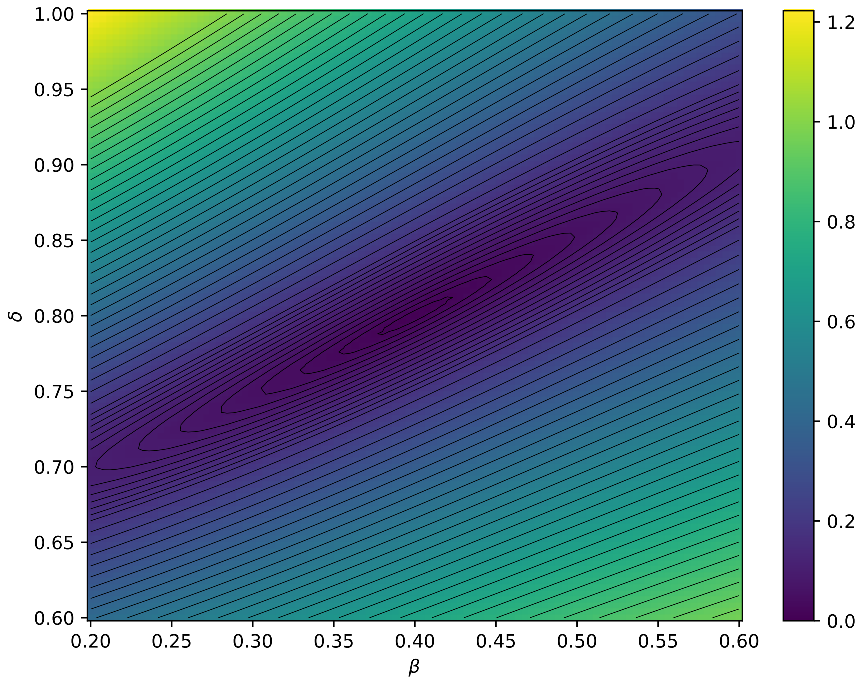

Figure 2: Distance criterion for the naive agent using the final periods of data

B.1 Simulation Results

Table 6: Estimation results for different sample sizes: Sophisticated agent, OLS estimator

	True value	$N = 50,000$	$N = 100,000$	$N = 1,000,000$
β	0.40	0.1727	0.2809	0.3718
		(0.7802)	(0.5290)	(0.1392)
δ	0.80	0.7838	0.7908	0.7947
		(0.1791)	(0.1209)	(0.0359)

Notes: Estimation results for the discount factors for different simulated sample sizes. Standard errors in parentheses.

Table 7: Estimation results for different sample sizes: Naive agent, MD estimator

	True value	$N = 50,000$	$N = 100,000$	$N = 1,000,000$
β	0.40	0.3953	0.4162	0.3991
		(0.1388)	(0.0940)	(0.0284)
δ	0.80	0.8008	0.8117	0.7995
		(0.0896)	(0.0634)	(0.0194)

Notes: Estimation results for the discount factors for different simulated sample sizes. Standard errors in parentheses.

Table 8: Estimation results for different sample sizes: Sophisticated agent, MLE (full model)

	True value	$N = 5,000$	$N = 10,000$	$N = 20,000$
β	0.40	0.3933	0.4060	0.4005
		(0.0954)	(0.0665)	(0.0457)
δ	0.80	0.7971	0.7990	0.7992
		(0.0381)	(0.0242)	(0.0166)
θ_1	2.50	2.5582	2.5092	2.5240
		(0.5054)	(0.2828)	(0.2234)
θ_2	0.70	0.8346	0.7231	0.7145
		(0.4965)	(0.2517)	(0.1602)
θ_3	0.00	0.0601	0.0058	0.0242
		(0.4849)	(0.2639)	(0.2135)
θ_4	1.00	1.1335	1.0236	1.0146
		(0.4870)	(0.2437)	(0.1550)

Notes: Estimation results for the discount factors for different simulated sample sizes. Standard errors in parentheses.

Table 9: Estimation results for different sample sizes: Naive agent, MLE (full model)

	True value	$N = 5,000$	$N = 10,000$	$N = 20,000$
β	0.40	0.4127	0.3866	0.4009
		(0.1197)	(0.0786)	(0.0624)
δ	0.80	0.8061	0.7967	0.7985
		(0.0537)	(0.0335)	(0.0288)
θ_1	2.50	2.6290	2.6493	2.5306
		(0.9243)	(0.6304)	(0.3692)
θ_2	0.70	0.7943	0.8002	0.7359
		(0.5419)	(0.3564)	(0.2307)
θ_3	0.00	0.1089	0.1515	0.0282
		(0.8887)	(0.6081)	(0.3669)
θ_4	1.00	1.0968	1.0996	1.0364
		(0.5319)	(0.3501)	(0.2266)

Notes: Estimation results for the discount factors for different simulated sample sizes. Standard errors in parentheses.

Table 10: Estimation results for different sample sizes: Sophisticated agent, MLE (full model with normalization)

	True value	$N = 5,000$	$N = 10,000$	$N = 20,000$
β	0.40	0.9714	0.9755	0.9804
		(6.7e-02)	(6.7e-02)	(5.2e-02)
δ	0.80	1.0000	1.0000	1.0000
		(2.4e-07)	(3.3e-14)	(1.5e-09)
θ_3	0.00	-1.1770	-1.2240	-1.1951
		(3.9e-01)	(3.8e-01)	(3.5e-01)
θ_4	1.00	0.2280	0.2363	0.2299
		(6.3e-02)	(6.0e-02)	(5.7e-02)

Notes: Estimation results for the discount factors for different simulated sample sizes. Standard errors in parentheses.

In the following, we provide graphical illustrations of the estimated CCPs for our different simulation settings.

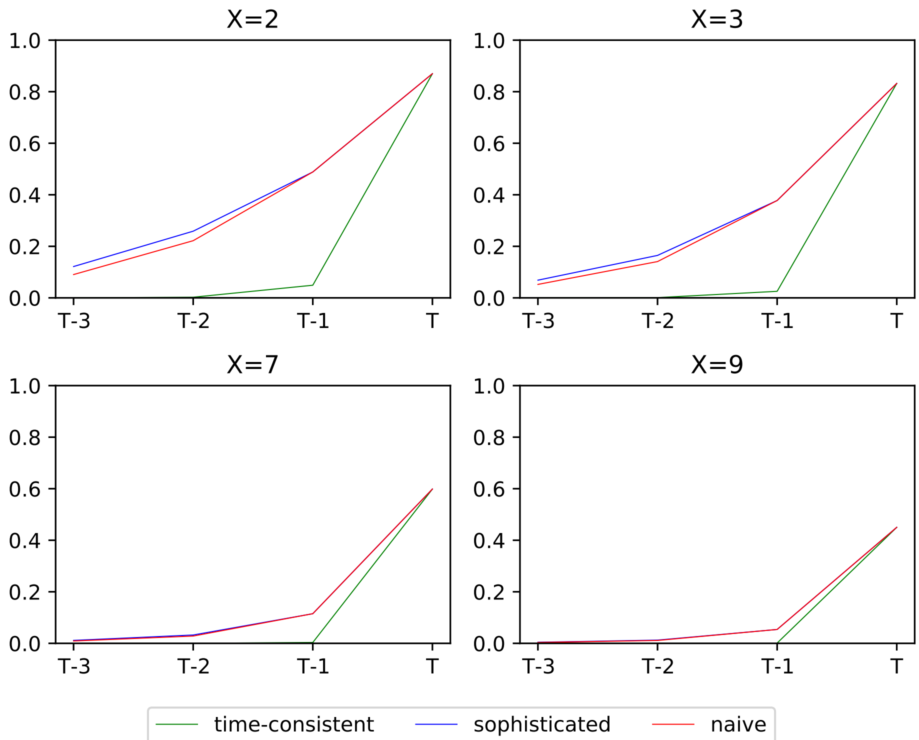

Figure 3: Adoption rate for different types of agents

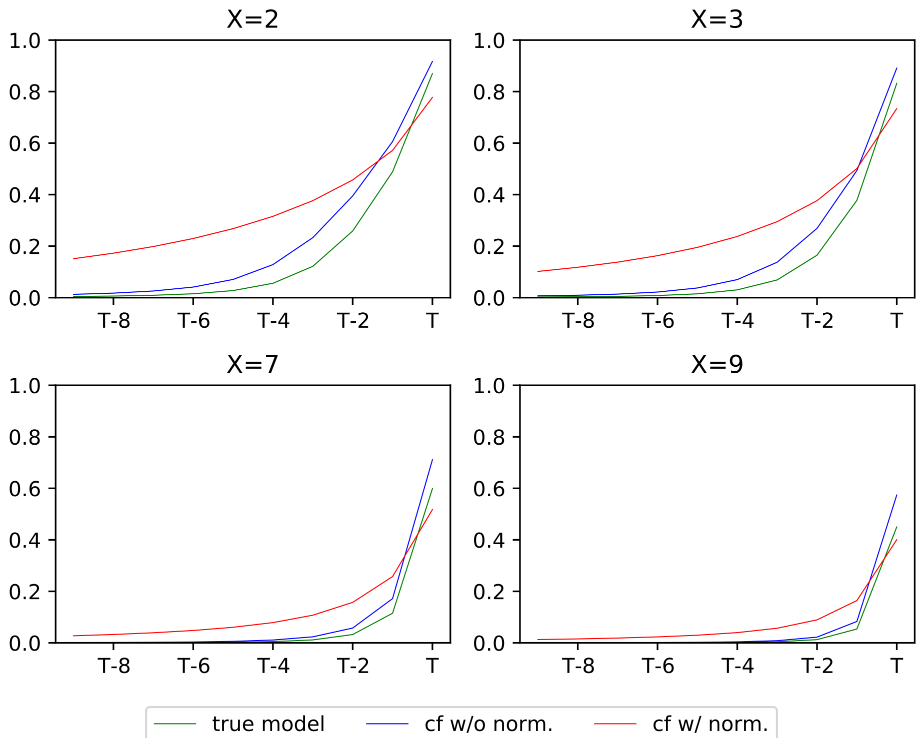

Figure 4: Counterfactual adoption rate (normalization on adoption)