CAEPR
Working Paper
#2022-010

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

Chao Wang
Indiana University, Department of Economics

Stefan Weiergraeber
Indiana University, Department of Economics

Ruli Xiao
Indiana University, Department of Economics

June 2, 2022

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

Chao Wang* Stefan Weiergraeber* Ruli Xiao*

June 2, 2022

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, Aprajit Mahajan, Brian McManus, and Christian Michel 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), labor (decisions to continue or drop out of school), 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]

Hyperbolically discounting agents can be either *sophisticated* or *naive*. The former are aware of their present-bias. The latter believe that they will not suffer from present-bias in the future. 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. 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.

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 model setup of flexible intertemporal decision making 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, in particular, present-biased, 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.[2] 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 standard DDCMs with exponential discounting, the discount factor is non-parametrically unidentified, see, for example, Rust (1994) and Magnac and Thesmar (2002). 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.

In general, 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. An additional problem of hyperbolic discounting models occurs when naive agents are considered. Because the current beliefs of a naive agent do not coincide with her actual actions in the future, it is generally impossible to recover these

[2] 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.

beliefs from the data. Therefore, incorporating hyperbolic discounting into DDCMs is not a trivial extension of the existing literature.

The literature on the identification of hyperbolic discount factors in finite horizon DDCMs is relatively nascent. For a finite horizon setup Abbring, O. Daljord, and Iskhakov (2019) provides 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.

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*.[3] 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. 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).

In an exponential discounting framework, one can write the difference in the (logged) choice probabilities associated with actions j and K (the terminating action) in period t as a function of the flow utility differences, the discount factor, and the continuation values associated with action j in period $t+1$. Differencing the choice probability contrasts across two periods will reveal the difference in the continuation value associated with action j across two periods, which is a simple function of 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. As a next step, one can always represent the continuation values as the flow utilities associated with the terminating action plus an adjustment term, which depends on the choice probabilities associated with the terminating action and the distribution of the unobserved error term.[4] Therefore, we can write the contrasts in the continuation value as contrasts of the adjustment terms, which allows us to identify the discount factor by looking at how the choice probability contrasts respond to contrasts in the continuation values.

[3] 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).

[4] See, for example, Arcidiacono and Miller (2011). If the error terms follow a type-1 extreme value distribution, the adjustment term is $m_K = -log(P_{Kt}) + \gamma$.

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.

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. Here, 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 for 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.

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 et al. (2014), and Kalouptsidi, Scott, and Souza-Rodrigues (2015).

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.[5]

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

[5]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 and consumer exits the market once she adopts the solar panel and thus 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.

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 general DDCMs with hyperbolic discounting.[6] 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 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, under their assumptions, 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 (2020) and Mahajan, Michel, and Tarozzi (2020), who derive the identification based on the 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.

[6]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 (2020) suggest some improvements over Fang and Wang (2015). In a similar spirit, Chan (2017) estimates a DDCM with hyperbolic discounting factor to analyze the welfare dependence.

Levy and Schiraldi (2021) 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. First, they study an infinite horizon setting, while we focus on a finite horizon model. Second, and 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 but with a certain delay, 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. In contrast, our identification strategy purely relies on the presence of a terminating action. Lastly, 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, i.e., discount factors and flow utilities for both the sophisticated and the naive agent. Section 4 illustrates that our proposed estimators perform well in simulations. Section 5 concludes.

2 General Hyperbolic Discounting: Model Setup

Consider a dynamic discrete choice model with decision periods indexed by $t = 1, ..., T$, where T can be finite or infinite. 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

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

where β captures the individual's present-bias, and δ 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, which depends on the future utility generated by the future self. Consequently, a forward-looking agent needs to predict how she will behave in the future, which depends on how the future self discounts her future utility. 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 therefore maximizes the life-time utility characterized by Equation 1 with $\beta = 1$.

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 (and forgo consumption) 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. The current self has a correct perception about her future self. Consequently, there is no conflict between the *current perceived* future self and the actual future self. In contrast, a naive agent thinks that it is hard to invest only today. She believes that it will be easier to make the investment in the future, because she (wrongly) believes that the discount rate in the future is only δ compared to $\beta\delta$ today. In the next period, however, the naive agent faces the same thought process again. Thus, a naive agent tends to postpone unpleasant actions, because she believes that her future self will put more weight on her future utilities, even though the future self in the future will behave exactly in the same way as the agent's 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. The *perceived* future self is time-consistent and maximizes the life-time utility while the actual future self follows the the same present-biased preferences as the current self.

The dynamic process can be summarized as follows. In each period t, the agent chooses an action k from a choice set $\mathcal{D} = \{1, 2, ..., K\}$. Prior to making the choice, the agent observes the state variables $x_t \in \mathcal{X} = \{x_1, \cdots, x_J\}$ and $\epsilon_t = \{\epsilon_{1,t}, ..., \epsilon_{K,t}\}$, 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)*

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

The state variables x_t are assumed to have finite support $\mathcal{X}$ and follow a stationary Markov process controlled by 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 from 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)*

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

Before proceeding to characterize the agent's optimization decision, we introduce some notation. The *current choice-specific value function* in period t is given by

$$w_{k,t}(x_t; \boldsymbol{\sigma}_{t+1}) = u_{k,t}(x_t) + \beta\delta \sum_{x_{t+1}} v_{t+1}(x_{t+1}; \boldsymbol{\sigma}_{t+1}) Q_k(x_{t+1}|x_t), \tag{4}$$

where $v_{t+1}(x_{t+1}; \boldsymbol{\sigma}_{t+1})$ is the perceived long-run value function, defined as

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

where $V_{t+1}(x_{t+2}, \epsilon_{t+2}; \boldsymbol{\sigma}_{t+2})$ is the value function under the perceived future self's strategy profile $\boldsymbol{\sigma}_{t+2} \equiv \{\sigma_\tau\}_{\tau=t+2}^T$. 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. Such a perceived long-run value function represents how an agent in the current period views the future after incorporating predictions about this own future behavior. 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 $\boldsymbol{\sigma}^*$ such that each σ_t^* is a best response to her perceived future strategy profile $\boldsymbol{\sigma}_{t+1}^*$, so that

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

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

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

where the mapping Φ_k depends on the distribution of the shocks.

In the infinite horizon framework, the existence of the equilibrium CCPs requires that the flow utility is time-invariant so that the long-run value function is a fixed point in Equation 8 given the CCPs. Consequently, the optimal behavior of the agent can be characterized by Equations (4), (5), and (7). By Brouwer's fixed point theorem, a perception perfect strategy exists.[7] In the finite horizon setting, 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. Her perception of the future self's strategies is consistent with the strategy actually chosen by the future self. That is, $\boldsymbol{\sigma}_{t+1}^*$ is consistent with the future equilibrium CCPs $p_{k,\tau}(x_\tau)$, where $\tau>t$. Similar to Arcidiacono and Miller (2011, Lemma 1), the perceived long-run value function $v_{t+1}(x_{t+1};\boldsymbol{\sigma}_{t+1}^*)$ for a sophisticated agent can be written as

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

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 $m_K(p(x))=\mathbb{E}_\epsilon max_k\left[w_k(x)-w_K(x)+\epsilon_k\right]$ is determined by the distribution of the ϵ-shocks. When ϵ_t follows an extreme value distribution, $m_K(p(x))=\gamma-logp_K(x)$, where γ is the Euler constant.

[7]The existence of perception perfect strategy for the same kind of stochastic games has been shown by (Peeters, 2004).

Naive agent. The key difference between naive agents and sophisticated agents is that the naive agent's actual behavior in the future diverges from the optimal strategies that the current self perceives today. To formalize the decision process of the current self, we introduce the *choice-specific value function of the next period self as perceived by the current self* $z_{k,t+1}(x)$

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

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

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

The *choice-specific value function* of the current self $w_{k,t}(x)$ is defined as

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

There are two differences between $w_{k,t}(x)$ and $z_{k,t+1}(x)$. First, in $w_{k,t}(x)$ the accumulated future payoff is discounted by $\beta\delta$, while in $z_{k,t+1}(x)$, the total future payoff is discounted by δ. Second, $w_{k,t}(x)$ represents how the current self actually evaluates the deterministic component of the payoff from choosing k, while $z_{k,t+1}(x)$ represents how the current self believes how her next-period self will evaluate the payoff from choosing k. $w_{k,t}(x)$ will determine the current self's optimal choice, but $z_{k,t+1}(x)$ governs the perception of the current self regarding the choices of her future selves. Finally, we define the future CCP perceived by the current self $\tilde{p}_{k_t,t}(x_t)$ as

$$\begin{aligned} \tilde{p}_{k,t+1}(x_{t+1}) &\equiv Pr[\tilde{\sigma}_{t+1}(x_{t+1}, \epsilon_{t+1}) = k] \qquad (12) \\ &= Pr[z_{k,t+1}(x_{t+1}; \tilde{\boldsymbol{\sigma}}_{t+2}) + \epsilon_{k,t+1} \geq z_{j,t+1}(x_{t+1}; \tilde{\boldsymbol{\sigma}}_{t+2}) + \epsilon_{j,t+1}, \quad \forall j \neq k]. \end{aligned}$$

This allows us to write the perceived long-run value function $v_{t+1}(x_{t+1}; \tilde{\boldsymbol{\sigma}}_{t+1})$ for a naive agent as

$$\begin{aligned} v_{t+1}(x_{t+1}; \tilde{\boldsymbol{\sigma}}_{t+1}) &= \mathbb{E}_{\epsilon_{t+1}} \{z_{\tilde{\boldsymbol{\sigma}}_{t+1},t+1}(x_{t+1}, \tilde{\boldsymbol{\sigma}}_{t+2})\} \\ &= \mathbb{E}_{\epsilon_{t+1}} \max_{k \in \mathcal{D}} [z_{k,t+1}(x_{t+1}, \tilde{\boldsymbol{\sigma}}_{t+2})] \\ &= m_K(\tilde{p}_{t+1}(x_{t+1})) + z_{K,t+1}(x_{t+1}; \tilde{\boldsymbol{\sigma}}_{t+2}). \qquad (13) \end{aligned}$$

Consequently, the equilibrium CCPs for a naive agent are characterized by Equations 7, 9, and 13. Note that the perceived CCPs are not observed in the data, because of the inconsistency between the actual and the currently perceived future self. However, the perceived instead of the actual CCPs in the future affect the equilibrium CCPs of the current self. This greatly complicates the identification since the perceived future self's CCPs are not observed in the data.

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.[8] 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, we rely on several assumptions. First of all, our key assumption is that there is some terminating action available to the agent. We denote the terminating action by K. 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:[9]

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

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

[8]We illustrate which parts of our identification arguments can be extended to an infinite-horizon setting in Appendix B.

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

over the future uncertainty as captured by the state transition matrix $\boldsymbol{Q}_{\boldsymbol{k}}$, 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, $\boldsymbol{u}_{k,t} - \boldsymbol{u}_{K,t}$ 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 3 *(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 preference 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 widely studied in the existing literature, see Mahajan, Michel, and Tarozzi (2020). 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, which is well known to affect counterfactual simulations, see Kalouptsidi, Scott, and Souza-Rodrigues (2015).

First of all, in the final period, the optimal decision is the same regardless of the agent's time preference because there is no future to discount. 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 4 that imposing such a normalization on the flow utility can substantially affect counterfactual simulations.

For now, we assume that the utility contrasts are known[10] 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.[11] 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., $\tilde{\boldsymbol{p}}_T = \boldsymbol{p}_T$. 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

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

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

CCPs in period $T-1$ as

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

where $\boldsymbol{v}_T(\tilde{\boldsymbol{p}}_T)$ 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 $\tilde{\boldsymbol{p}}_T = \boldsymbol{p}_T$ 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 developed in the seminal work of Hotz and Miller (1993).

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.

Assumption 4 *(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 the difference results in column sum being 0.

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 $\boldsymbol{u}_k - \boldsymbol{u}_K$ is known. That is,

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

where $\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-1}) = \boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-1}) - \boldsymbol{\phi}_{kK}(\boldsymbol{p}_T)$. 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.

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 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, we can identify and estimate the agent type-specific CCPs using insights from the measurement error or finite mixture literature, see, for example, Hu (2008). We can then follow the identification results developed in this paper using the type-specific CCPs to identify the flow utilities and the discount factors for each type.

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 $\sum_k Q_k(x_T = j|x_{T-1} = i)p_{k,T-1}(x_{T-1} = i)$. Because almost all components in Equation 20 are identified from data on the last two periods, the log odds ratio in period $T-2$ only contains the two discount factors as unknowns.

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 $T-1$, because there is no future after the period following $T-1$. 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,

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

which 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 $A \equiv \boldsymbol{Q}_k(log(\boldsymbol{p}_{KT}) - log(\boldsymbol{p}_{KT-1}) - \bar{\boldsymbol{Q}}_{T-1}\boldsymbol{Q}_k^{-1}\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_T))$ 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 5 *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. We summarize the above discussion as

Proposition 1 *If Assumptions 1 to 5 are satisfied and if there is (at least) one terminating action K, 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. Stationary utility function 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 four periods in Proposition 4 in the later section. 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 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 (2015). We provide an illustration of the generated bias in our simulation in Section 4.

3.1.2 Naive agents

In this section we provide identification results for the hyperbolic discounting parameters of a 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

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

where $\tilde{\boldsymbol{p}}_{T-1}$ 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., $\tilde{\boldsymbol{p}}_{T-1} \neq \boldsymbol{p}_{T-1}$. 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 $\delta\beta$. Such an inconsistency between perception and actual behavior greatly complicates the identification.[12]

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 selfs to recover the current self's beliefs, we use the current self's choices to recover her perceptions about the future. Specifically, from the actual log odds ratios of the $T-2$-self observed in the data, we can recover her perception of the behavior of her $T-1$-self and T-self, which is captured by the perceived ex ante value $\boldsymbol{v}_{T-1}(\tilde{\boldsymbol{p}}_{T-1})$. That is, from Equation 23 and with the full rank condition on the transition matrix, imposed by Assumption 4, we

[12]If the agent is sophisticated, we always have $\tilde{\boldsymbol{p}}_{T-1} = \boldsymbol{p}_{T-1}$, which is identifiable from the data directly. In this case, combining Equations 18 and 23 allows us to identify $\boldsymbol{u}_k$ and $\beta\delta$.

have

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

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 as perceived by the naive $T-2$-self can be represented as[13]

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

where $\boldsymbol{v}_T(\tilde{\boldsymbol{p}}_T)$ captures how the $T-2$-self believes her $T-1$-self to think about period T. 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 (McFadden, 1977):

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

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 function identified from Equation 25, we have

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

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

[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 $T-1$.

Assumption 6 *The gradient of the J restrictions has a rank of 2 at the true parameter values, where the restrictions are defined as*

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

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.

Proposition 2 *If Assumptions 1 -4 and 6 holds 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.*

To the best of our knowledge, we are the first to study point identification of the naive agent's time-preference in the DDC 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. In this subsection, we extend our identification strategy to settings in which the final periods are not observed in the data.

3.2.1 Sophisticated Agent

Our starting point is the log odds ratio vector in period t, which is obtained by stacking the log odds ratios from Equation (15) for each state value x:

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

where $\boldsymbol{\phi}_{kK}(\boldsymbol{p}_t)$ is a $J \times 1$ vector. Given this expression, we can difference the CCPs in two consecutive periods so that the log odds contrast becomes

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

where $t \geq T-1$ and $\boldsymbol{v}_{T+1} = 0$ if T is the final period. 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.[14] 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.

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

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 4 holds, we have

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

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 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:

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

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 7 *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,*

2. *Assumptions 1 -4 and 7 hold, and*

3. *there is a terminating action K,*

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, has J restrictions, and 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 8 *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 4, 7, and 8 are satisfied and if there is (at least) one terminating action K, 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.*

Note that when we observe agent's 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 $\boldsymbol{u}_{\boldsymbol{K}}$ is known; Abbring, O. Daljord, and Iskhakov (2019) assume that $\boldsymbol{u}_{\boldsymbol{K}} = 0$.

3.2.2 Naive Agent

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

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

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 $\tilde{\boldsymbol{p}}_\tau$ 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

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

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 $\boldsymbol{m}_K(\tilde{\boldsymbol{p}}_t)$.

The key difficulty for identification of the model with naive agents is that we cannot directly observe how the naive agent believes her future self to behave. Therefore, we cannot recover $\tilde{p}_{t+1}$ from the data directly 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 data on the final three periods.

4 Monte Carlo Simulations

In this section, we analyze a stylized model setup to illustrate that our proposed estimators work well in simulations.[15] First, we introduce the model setup and its parametrization.[16] Next, we propose to estimate the two discount factors closely following the identification argument presented in Section 3. The payoff primitives can be estimated nonparametrically after the estimation of the two discount factors. We illustrate the finite sample properties of the discount factor estimators. Not suprisingly, these estiamtors requires a large sample size in order to estimate the parameters precisely. Therefore, we propose to estimate the full model, i.e., discount factors and flow utilities jointly, using maximum likelihood to improve efficiency. 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. We demonstrate how the rank conditions required by our identification assumptions can be checked in practice in the Appendix.

4.1 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.

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

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

We specify the flow utility as

$$u(x,a) = \begin{cases} \theta_1 + \theta_2 x & \text{if } a_i = 1 \\ \theta_3 + \theta_4 x & \text{if } a_i = 0, \end{cases} \tag{36}$$

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

$$Q_0(x'|x) = \begin{bmatrix} 0.4800 & 0.2400 & 0.1600 & 0.1200 \\ 0.2143 & 0.4286 & 0.2143 & 0.1429 \\ 0.1429 & 0.2143 & 0.4286 & 0.2143 \\ 0.1200 & 0.1600 & 0.2400 & 0.4800 \end{bmatrix} \tag{37}$$

Given the model primitives, we can solve for the equilibrium CCPs. Figure 3 in Appendix C.1 presents the adoption rates for different types of agents (time-consistent, sophisticated, and naive) in different states.[17] As x becomes larger, the adoption rate decreases since θ_2 is smaller than θ_4, so that utility from waiting becomes relatively larger. For the same x, the adoption rate becomes larger as time goes by. 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.

4.2 Estimation of the Discount Factors only

We first estimate the two discount factors only following the identification restrictions discussed in Section 3. Note that our paper mostly focuses on identification and not estimation. In order to keep the illustration concise, we focus on the setting in which the last three periods are observed in the data. We can estimate the two discount factors via an ordinary least square (OLS) estimator for sophisticated agents and a minimum distance estimator for naive agents. We also illustrate the performances of the two estimators in finite samples.

Suppose we observe data $\{a_{it}, x_{it}\}_{i=1,\ldots,N,t=T-2,T-1,T}$. We first estimate the equilibrium CCPs for adopting the solar system in period t via a simple frequency estimator. That is,

$$\hat{p}_t(a_t = 1|x_t = x) = \frac{\sum_i I(a_{it} = 1, x_{it} = x)}{\sum_i I(x_{it} = x)}, \; t = T-2, T-1, T. \tag{38}$$

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

The transition matrix of the state x is estimated similarly using the simple frequency estimator. We now describe how to estimate the two discount factors for both sophisticated and naive agents, respectively.

Sophisticated agent. For sophisticated agents, we can estimate the two discount factors using OLS following the identification condition in Equation 22. Specifically,

$$\begin{bmatrix} \hat{\delta}\hat{\beta} \\ \hat{\delta} \end{bmatrix} = \left(\Omega'(\hat{\boldsymbol{p}}_{T-2}, \hat{\boldsymbol{p}}_{T-1}, \hat{\boldsymbol{p}}_T)\Omega(\hat{\boldsymbol{p}}_{T-2}, \hat{\boldsymbol{p}}_{T-1}, \hat{\boldsymbol{p}}_T)\right)^{-1} \Omega'(\hat{\boldsymbol{p}}_{T-2}, \hat{\boldsymbol{p}}_{T-1}, \hat{\boldsymbol{p}}_T)\Delta\boldsymbol{\phi}_{kK}(\hat{\boldsymbol{p}}_{T-1}). \quad (39)$$

We simulate the data based on the DGP presented in Section 4.1 and summarize the estimation results in Table 1 for sample sizes of 50,000, 100,000, and 1,000,000. Throughout, our results are based on 100 simulation runs. 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 the sample sizes are reported by Abbring and Ø. Daljord (2020).

Table 1: OLS estimation results for the sophisticated agent and different sample sizes

	True value	$N = 50,000$	$N = 100,000$	$N = 1,000,000$
β	0.40	0.2757	0.3863	0.3757
		(0.6649)	(0.3996)	(0.1337)
δ	0.80	0.7988	0.8097	0.7949
		(0.1559)	(0.1049)	(0.0349)

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

Naive agent. Next, we estimate the two discount factors for naive agents using a minimum distance estimator, following the identification restrictions in Equation 27. That is,

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

where $||\cdot||$ denote the L-2 norm. We simulate the data using the true CCPs for sample sizes of 50,000, 100,000, and 1,000,000, and estimate the two discount factors. For each sample size, we replicate the estimation 100 times. Table 2 presents the mean and standard deviation from these estimations. Most notably, both discount factors are estimated precisely for all of our sample sizes.

Table 2: Estimation results for the naive agent and different sample sizes

	True value	$N = 50,000$	$N = 100,000$	$N = 1,000,000$
β	0.40	0.4093	0.4082	0.3984
		(0.1489)	(0.1054)	(0.0287)
δ	0.80	0.8094	0.8085	0.7993
		(0.0908)	(0.0677)	(0.0192)

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

4.3 Estimation of the Full Model

In practice, estimating both the discount factors and the flow utilities jointly non-parametrically is often too demanding on the data. Instead, researchers typically estimate the payoff function parametrically. In this subsection, we illustrate the performance of a typical maximum likelihood estimator that estimates the discount factors and the parameterized flow utility jointly. Let θ denote all parameters, i.e., $\theta \equiv \{\beta, \delta, \theta_1^u, \theta_0^u\}$, where θ_1^u and θ_0^u captures the parameters in the flow utility associated with action 1 and 0, respectively. If we do not impose any restrictions on the flow utility, the log-likelihood function is standard and given by

$$\hat{\theta} = \arg max_\theta \sum_i \sum_t log\left(p_t(a_{it}|x_{it};\theta)\right). \quad (40)$$

We estimate the model for sample sizes of 5,000, 10,000, and 20,000 and run each simulation 100 times for both the sophisticated and the naive agents. Table 3 and 4 present the results for the sophisticated agent and the naive agent, respectively. For both the sophisticated and the naive agent 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 model under a normalization condition on the flow

Table 3: Sophisticated agent: Estimation results for full model using MLE for different sample sizes

	True value	$N = 5,000$	$N = 10,000$	$N = 20,000$
β	0.40	0.4000	0.4052	0.4021
		(0.1057)	(0.0734)	(0.0533)
δ	0.80	0.7982	0.8003	0.7970
		(0.0345)	(0.0251)	(0.0190)
θ_1	2.50	2.5044	2.5178	2.5113
		(0.3508)	(0.2818)	(0.2241)
θ_2	0.70	0.8306	0.7389	0.7253
		(0.4800)	(0.3625)	(0.1980)
θ_3	0.00	-0.0014	0.0146	0.0120
		(0.3153)	(0.2698)	(0.2231)
θ_4	1.00	1.1299	1.0396	1.0257
		(0.4668)	(0.3551)	(0.1929)

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

Table 4: Naive agent: Estimation results for full model using MLE

for different sample sizes

	True value	$N = 5,000$	$N = 10,000$	$N = 20,000$
β	0.40	0.4301	0.4053	0.4160
		(0.1143)	(0.0844)	(0.0555)
δ	0.80	0.8155	0.8032	0.8113
		(0.0530)	(0.0365)	(0.0300)
θ_1	2.50	2.4957	2.4846	2.3548
		(0.7206)	(0.5994)	(0.4438)
θ_2	0.70	0.7159	0.7588	0.6863
		(0.4581)	(0.3484)	(0.1808)
θ_3	0.00	-0.0254	-0.0083	-0.1398
		(0.6937)	(0.5817)	(0.4215)
θ_4	1.00	1.0190	1.0574	0.9854
		(0.4504)	(0.3426)	(0.1781)

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

utility. In this case, the maximum likelihood estimator is modified to incorporate the normalization as the follows:

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

For this exercise, we focus on the setting in which we observe the final period in the data and the sophisticated agent and we estimate the technology adoption model described in Section 4.1. The estimation results when normalizing the utility associated with the adoption action are displayed in Table 5.[18] In our maximum likelihood estimation, that estimates utilities and discount factors jointly, the normalization of the flow utility also leads to biased estimates of the discount factors. Most notably, both discount factors are substantially overestimated and the exponential discount factor often hits the upper bound of 1.

Table 5: Sophisticated agent: Estimation results for full model with normalization for different sample sizes

	True value	$N = 5,000$	$N = 10,000$	$N = 20,000$
β	0.40	0.9723	0.9747	0.9799
		(6.9e-02)	(6.5e-02)	(4.9e-02)
δ	0.80	1.0000	1.0000	1.0000
		(2.4e-09)	(6.2e-05)	(8.8e-08)
θ_3	0.00	-1.2015	-1.2059	-1.1883
		(3.8e-01)	(3.9e-01)	(3.4e-01)
θ_4	1.00	0.2324	0.2331	0.2286
		(6.0e-02)	(6.2e-02)	(5.7e-02)

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

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 normalization on non-adoption case, or θ_3 is increased by 0.5 in normalization on adoption case. We plot the counterfactual CCPs with and without normalization in Figure 4 in

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

Appendix C.1.[19] We can see that that under a subsidy on adopting the model without normalization increases the adoption rate slightly compared to the status quo, while the model with normalization raises 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 (2015).

5 Conclusion

In this paper, we study the identification of dynamic discrete choice models (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 as well as their flow utilities. Our identification strategy exploits the recursive structure of a 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 as long as we observe four 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 is likely to bias counterfactual simulations, see, for example, Kalouptsidi, Scott, and Souza-Rodrigues (2015).

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

[19] Here we plot the simulated CCPs for ten periods to see the long-run dynamic effects. The graphs here based on sample size of 5,000.

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 (May 2020). "a Comment on "Estimating Dynamic Discrete Choice Models With Hyperbolic Discounting" By Hanming Fang and Yang Wang". In: *International Economic Review* 61.2, pp. 565–571. ISSN: 14682354. DOI: 10.1111/iere.12434. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/iere.12434.

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

Aguirregabiria, Victor et al. (2014). "Identification and counterfactuals in dynamic models of market entry and exit and Wisconsin-Madison, for helpful comments and suggestions". In: 12, pp. 267–304. DOI: 10.1007/s11129-014-9147-5.

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

Arcidiacono, Peter and Robert A. Miller (Nov. 2011). "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity". In: *Econometrica* 79.6, pp. 1823–1867. ISSN: 0012-9682. DOI: 10.3982/ecta7743. URL: https://onlinelibrary.wiley.com/doi/full/10.3982/ECTA7743%20https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA7743%20https://onlinelibrary.wiley.com/doi/10.3982/ECTA7743.

Bajari, Patrick et al. (2016). "Identification and semiparametric estimation of a finite horizon dynamic discrete choice model with a terminating action". In: *Quantitative Marketing and Economics* 14.4, pp. 271–323. ISSN: 1573711X. DOI: 10.1007/s11129-016-9176-3.

Blevins, Jason R. et al. (2020). "a Dynamic Discrete Choice Model of Reverse Mortgage Borrower Behavior". In: *International Economic Review* 00.00. ISSN: 14682354. DOI: 10.1111/iere.12470.

Chan, Marc K. (Oct. 2017). "Welfare dependence and self-control: An empirical analysis". In: *Review of Economic Studies* 84.4, pp. 1379–1423. ISSN: 1467937X. DOI: 10.1093/restud/rdx011. URL: http://academic.oup.com/restud/article/84/4/1379/3091894/Welfare-Dependence-and-SelfControl-An-Empirical.

Colas, Mark, Sebastian Findeisen, and Dominik Sachs (Feb. 2021). "Optimal need-based financial aid". In: *Journal of Political Economy* 129.2, pp. 492–533. ISSN: 1537534X. DOI: 10.1086/711952/SUPPL_FILE/2017507DATA.ZIP. URL: https://www.journals.uchicago.edu/doi/abs/10.1086/711952.

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"". In: *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". In: *American Economic Review* 109.6, pp. 2137–2172. ISSN: 19447981. DOI: 10.1257/aer.20161343. URL: https://doi.org/10.1257/aer.20161343.

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

Fang, Hanming and Yang Wang (May 2015). "Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions". In: *International Economic Review* 56.2, pp. 565–596. ISSN: 14682354. DOI: 10.1111/iere.12115. URL: http://doi.wiley.com/10.1111/iere.12115.

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

Hotz, V. Joseph and Robert A. Miller (July 1993). "Conditional Choice Probabilities and the Estimation of Dynamic Models". In: *The Review of Economic Studies* 60.3, p. 497. ISSN: 00346527. DOI: 10.2307/2298122. URL: https://academic.oup.com/restud/article-lookup/doi/10.2307/2298122.

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

Kalouptsidi, Myrto (2014). "Time to build and fluctuations in bulk shipping". In: *American Economic Review* 104.2, pp. 564–608. ISSN: 00028282. DOI: 10.1257/aer.104.2.564. URL: http://dx.doi.org/10.1257/aer.104.2.564564.

Kalouptsidi, Myrto, Paul Scott, and Eduardo Souza-Rodrigues (Sept. 2015). "Identification of Counterfactuals in Dynamic Discrete Choice Models". In: *CEPR Discussion Paper*. DOI: 10.3386/w21527. URL: http://www.nber.org/papers/w21527.pdf.

Laibson, David (May 1997). "Golden eggs and hyperbolic discounting". In: *Quarterly Journal of Economics* 112.2, pp. 442–477. ISSN: 00335533. DOI: 10.1162/003355397555253. URL: https://academic.oup.com/qje/article/112/2/443/1870925.

Levy, Matthew and Pasquale Schiraldi (2021). "Identification of intertemporal preferences from choice set variation". In: *Available at SSRN 3834344*.

Magnac, Thierry and David Thesmar (2002). *Identifying dynamic discrete decision processes*. Tech. rep. 2, pp. 801–816.

Mahajan, Aprajit, Christian Michel, and Alessandro Tarozzi (2020). *Identification of Time-Inconsistent Models: The Case of Insecticide Treated Nets*. Tech. rep. DOI: 10.3386/w27198. URL: http://www.nber.org/papers/w27198.

McFadden, Daniel (1977). "Modelling the choice of residential location". In.

Norets, A. and X. Tang (2014). "Semiparametric inference in dynamic binary choice models". In: *Review of Economic Studies* 81.3, pp. 1229–1262. ISSN: 1467937X. DOI: 10.1093/restud/rdt050. URL: https://academic.oup.com/restud/article-abstract/81/3/1229/1599945.

O'Donoghue, Ted and Matthew Rabin (1999). "Doing it now or later". In: *American Economic Review* March, pp. 103–124. DOI: 10.2307/j.ctvcm4j8j.12.

Peeters, R (2004). "Stochastic Games with Hyperbolic Discounting". In: URL: https://core.ac.uk/download/pdf/231371786.pdf.

Rust, John (Jan. 1994). *Chapter 51 Structural estimation of markov decision processes*. DOI: 10.1016/S1573-4412(05)80020-0.

Appendices

The Appendix consists of three parts: the detailed proofs for the lemmas and theorems presented in the main text, the identification results extended to infinite horizon models, 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:

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

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:

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

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:

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

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

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

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 4, 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,

$$rank(D \times E) \geq rank(D) + rank(E) - J = J, \tag{44}$$

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 Identification of Infinite Horizon Framework

In this section, we investigate the identification of the sophisticated agent's preferences in an infinite horizon model. Notice that the variation of choice probabilities over time is the key source of identification when the flow utility is stationary. In the infinite horizon framework, there is no variation in choice probabilities over time due to the stationarity of the decision process. Consequently, we need to rely on other exclusion restrictions. Note that the existing literature mostly focusses on the finite horizon framework. The model setup is similar to that in Section 3.2, but we drop the time subscript because of to stationary.

As in the finite horizon framework, the identification of the naive agent's preferences is more challenging than for the sophisticated agent due to the inconsistency between the choices of future selfs as perceived by the current self and the actual counterpart. Therefore, we focus on the sophisticated agent framework. Similarly to the finite horizon framework, the log odds ratios in Equation (15) can be written in the following vector form:

$$\boldsymbol{\phi}_{kK}(p(x)) = w_k(x) - w_K(x) = u_k(x) - u_K(x) + \beta\delta\boldsymbol{Q}_k(x)\boldsymbol{v}, \tag{45}$$

from which we can see that the log odds ratios are jointly determined by the flow utility differences between action k and K, the compound discount rate $\delta\beta$, the transition matrix $\boldsymbol{Q}_k$, and the future *ex ante* value function $\boldsymbol{v}$. Now that the transition matrix can be identified from the data separately with the assumption that the agent holds rational expectation regarding future uncertainty.

To understand how the expected utility is determined, we plug the choice-specific value function $w_K(x) = u_K(x)$ because action K is a terminal action into Equation (8) and rewrite the *perceived long-run value function* in vector form (stacking over x):

$$\boldsymbol{v} = \boldsymbol{m}_K + \boldsymbol{u}_K + \delta(1-\beta)\bar{\boldsymbol{Q}}\boldsymbol{v}, \tag{46}$$

which characterizes the long-run value function as a fixed-point in the above mapping. Intuitively, one can recover the expected long-term utility by iterating Equation ()46) into the future and collecting all realized flow utilities. That is, we can directly express the expected value function as a function of the two discount factors and the flow utility u_K.

$$\boldsymbol{v} = \left(I - \delta(1-\beta)\bar{\boldsymbol{Q}}\right)^{-1}(\boldsymbol{m}_K + \boldsymbol{u}_K), \tag{47}$$

Furthermore, if we impose the standard normalization assumption, i.e., $u_K = 0$, one can represent the expected value as a function of the discount factors, see Abbring, O. Daljord, and Iskhakov (2019).

Plugging in the expression for the long-term value function as described by Equation (47), the log odds ratio becomes:

$$\boldsymbol{\phi}_{kK}(p(x)) = u_k(x) + \beta\delta \boldsymbol{Q}_k(x) \left(I - \delta(1-\beta)\bar{\boldsymbol{Q}}\right)^{-1} \boldsymbol{m}_K. \tag{48}$$

The log odds ratio of action k and K for state variable x depends on the utility associated with action k (since we impose the normalization condition), the future uncertainty captured by the transition vector $\boldsymbol{Q}_k(x)$[20], the compound discount factor $\beta\delta$, and the long-run value function, which is fully captured by the observed CCPs and the discount factors.

To control for the direct impact of the flow utility, exclusion restrictions are required, as described in the following assumption.

Assumption 9 *There exist two pairs of states $x_{a,1}$,$x_{a,2}$ and $x_{b,1}$,$x_{b,2}$, such that*

$$u_k(x_{a,1}) = u_k(x_{a,2}) \quad \textit{and} \quad u_k(x_{b,1}) = u_k(x_{b,2}),$$

where $x_{a,1} \neq x_{a,2}$ and $x_{b,1} \neq x_{b,2}$.

With such an exclusion restriction, we can get the bivariate system of equations in β and δ:

$$\begin{aligned} \boldsymbol{\phi}_{kK}(p(x_{a,1})) - \boldsymbol{\phi}_{kK}(p(x_{a,2})) &= \beta\delta \left[\boldsymbol{Q}_k(x_{a,1}) - \boldsymbol{Q}_k(x_{a,2})\right] \left(I - \delta(1-\beta)\bar{\boldsymbol{Q}}\right)^{-1} \boldsymbol{m}_K \\ \boldsymbol{\phi}_{kK}(p(x_{b,2})) - \boldsymbol{\phi}_{kK}(p(x_{b,2})) &= \beta\delta \left[\boldsymbol{Q}_k(x_{b,1}) - \boldsymbol{Q}_k(x_{b,2})\right] \left(I - \delta(1-\beta)\bar{\boldsymbol{Q}}\right)^{-1} \boldsymbol{m}_K \end{aligned} \tag{49}$$

This system of equations can provide bounds for the discount factors β and δ.

C 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. To simplify the illustration and to avoid finite sample error, we assume that we know the population CCPs and the state transition matrix. We discuss sophisticated and naive agents separately. Given that all our technical assumptions are rank conditions, we rely on the singular values[21] of the matrix to verify that all regularity conditions imposed by our assumptions are satisfied.

Sophisticated agent. Our first step is to verify the full rank condition on the state transition matrix Q_0 imposed by Assumption 4. Table 6 displays the singular values for matrix

[20] Since action K is terminating action, $Q_K = 0$.

[21] 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.

$\boldsymbol{Q}_0$. 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 4 is satisfied in our example.

SV1	SV2	SV3	SV4
1.0010	0.3923	0.2426	0.1821

Table 6: Singular values for $\boldsymbol{Q}_0$

Second, we verify the rank condition specified in Assumption 5, 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 7. 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 7: Ω and its singular values

Once the rank condition is satisfied, we can directly compute the two discount factors using Equation (22)

$$\begin{bmatrix} \delta\beta \\ \delta \end{bmatrix} = \left(\Omega'(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)\Omega(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)\right)^{-1} \Omega'(\boldsymbol{p}_{T-2}, \boldsymbol{p}_{T-1}, \boldsymbol{p}_T)\Delta\boldsymbol{\phi}_{kK}(\boldsymbol{p}_{T-1}). \quad (50)$$

The computation result is shown in Table 8. The computed discount factors are perfectly consistent with the true parameters.

$\delta\beta$	δ
0.3200	0.8000

Table 8: Verified discount factors

Third, we examine the condition imposed by Assumption 7. Table 9 presents the differences between the CCP of waiting (action 0) in different time periods.[22]. All values are nonzero, which means that the non-zero condition in Assumption 7 holds in our example and both matrices seem well-posed.

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

	$p_{0,T-3} - p_{0,T-2}$	$p_{0,T-2} - p_{0,T-1}$	$p_{0,T-1} - p_{0,T}$
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 9: $p_{k,t} - p_{k,t+1}$

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.[23] We present the associated results in Table 10. 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.

	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 10: true u_1 and computed u_1 from three consecutive periods of data

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 8.

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:

$$\Delta(\beta, \delta) \equiv |\Upsilon(\beta, \delta, \boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) - \Upsilon(\beta, \delta, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4)|, \tag{51}$$

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.[24]

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

[24]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

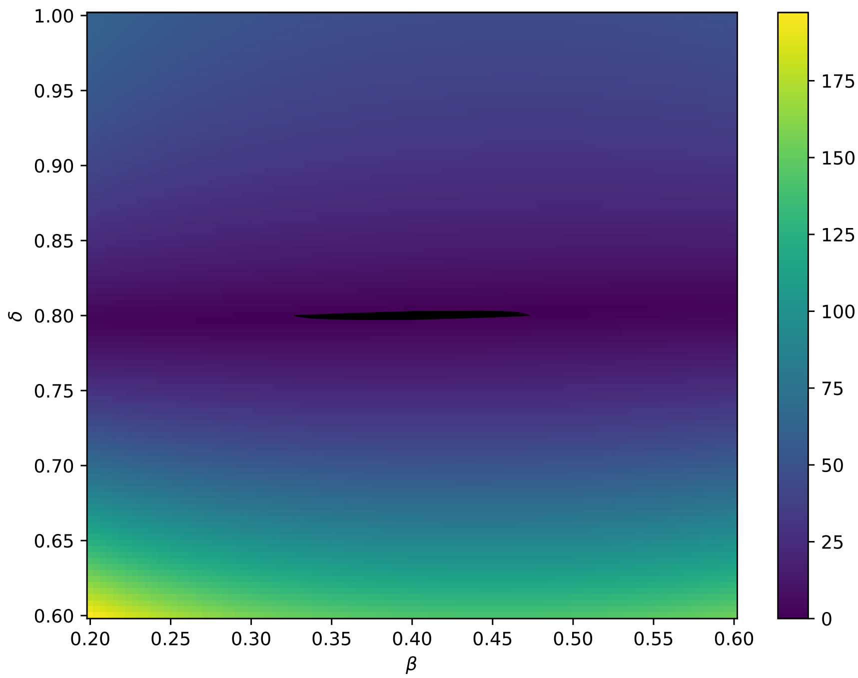

Figure 1: Distance criterion for sophisticated agent using four consecutive periods of data

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 6, which can be done analogously to the sophisticated agents 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.

C.1 Additional Counterfactual Results

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

between 0 and 1.

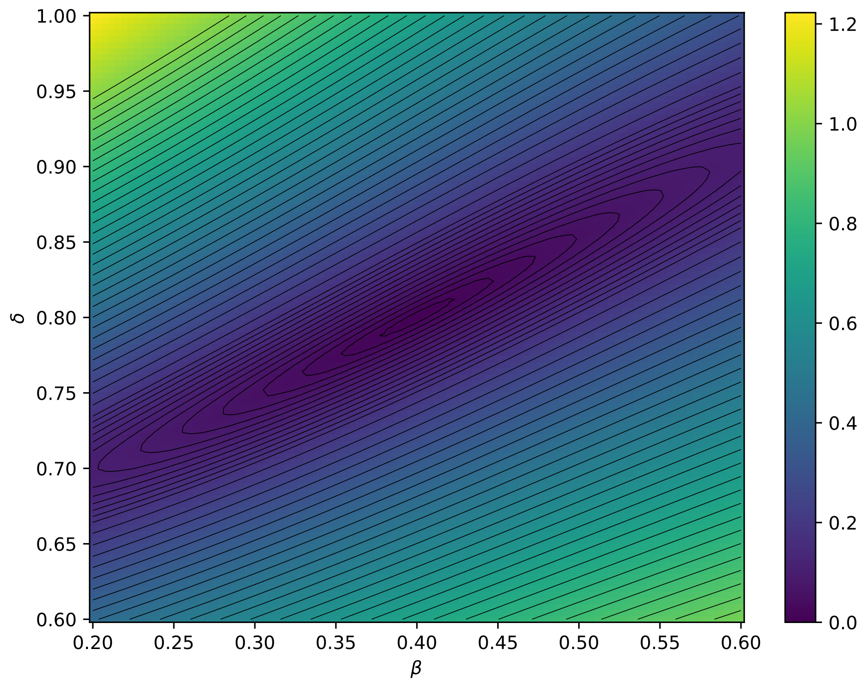

Figure 2: Distance criterion for the naive agent using the final periods of data

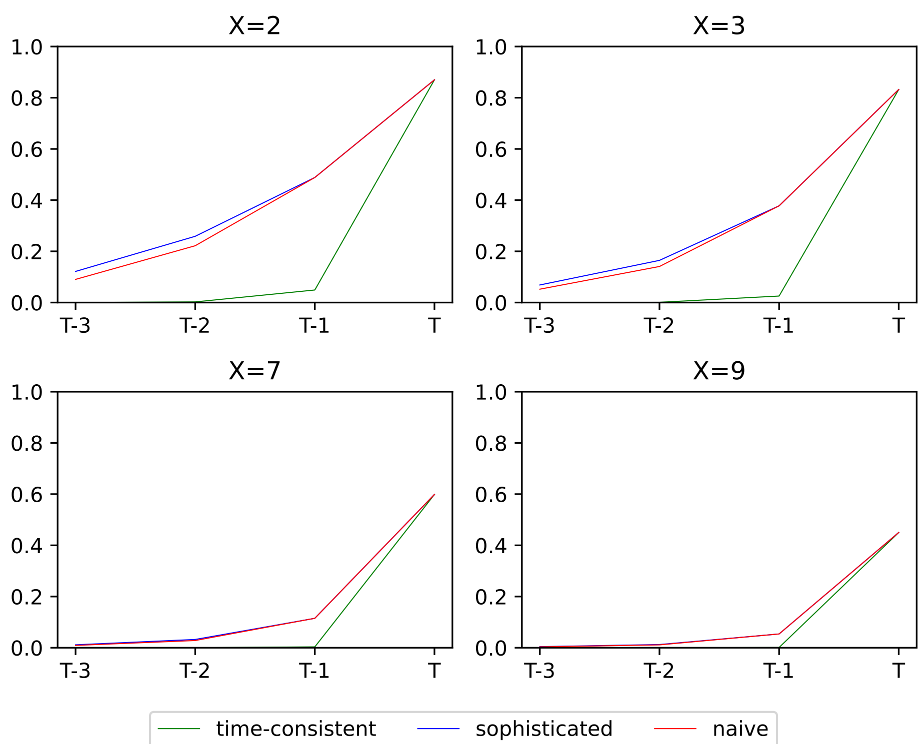

Figure 3: Adoption rate for different types of agents

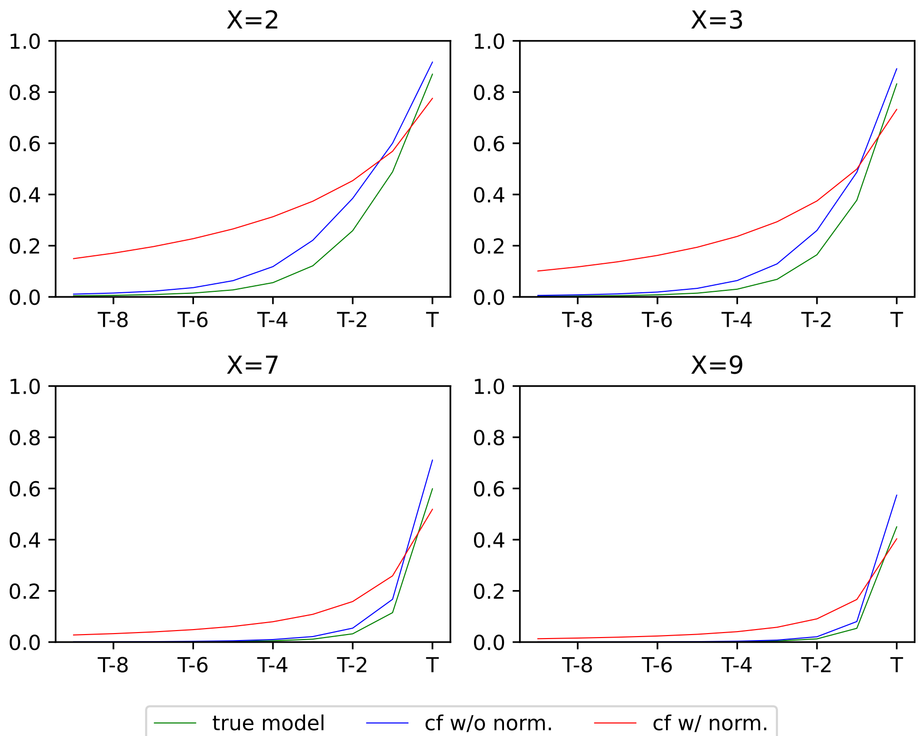

Figure 4: Counterfactual adoption rate (normalization on adoption)