Submitted to *Econometrica*

NEURAL FOUNDATIONS

QIAN QI
Peking University, No.5 Yiheyuan Road Haidian District, Beijing, P.R.China 100871

"Intelligent agents are rational, yet limited resources may lead to irrational behavior."

This paper introduces **Neural Differential Utility (NDU)**, a novel axiomatic framework integrating deep neural networks within a recursive utility paradigm. NDU offers a tractable yet more granular approach for analyzing the evolution of preferences within a **resource-constrained agent**. We demonstrate NDU's capability to endogenize nonlinear and time-varying preferences via representative examples, shedding light on the theoretical foundations that underpin the growing applications of deep learning in economic analysis. Five companion papers investigate the examples, Universal Approximation Theorem of NDU, and examine its implications for asset pricing.

KEYWORDS: Artificial intelligence, Endogenous Preferences, Resource Constraint, Deep Learning Theory, Mean-field Control, Recursive Utility.

"The intellect seeking after an integrated theory cannot rest content with the assumption that there exist two distinct fields totally independent of each other by their nature," ***Albert Einstein*** said in his Nobel lecture in 1923.

1. INTRODUCTION

A fundamental challenge in economics is understanding the formation and evolution of individual preferences. Traditional economic analysis frequently relies on the assumption of stable, exogenously determined preferences. While such models offer valuable insights into market behavior and policy design (e.g., Debreu

Qian QI: qiqian@pku.edu.cn

The initial idea commenced in November 2013, with an early preliminary draft finalized in June 2014. I would like to extend my gratitude to Darrell Duffie, Sarah Auster, for their insightful and very detailed discussions. I appreciate the valuable feedback from Zengjing Chen on an early version, which elucidate the key limitations of the methodologies employed in earlier drafts. I also thank Thomas Sargent for his comments and encouragements.

et al. (1954), Arrow (1963)), they struggle to explain phenomena where individual choices demonstrably evolve over time.[1] The Lucas critique (Lucas Jr (1976)) highlighted the importance of accounting for how policy changes themselves can reshape the behavior they are intended to influence, underscoring the need to consider the endogeneity of preferences. Subsequent work on microfoundations, such as (Lucas Jr (1978), Lucas and Sargent (1978), Prescott and Mehra (1980)), and most recently (e.g., Sargent (2024a,b)), further emphasized the need to understand the micro-level forces driving individual decision-making.

A crucial aspect of this understanding is recognizing that preferences are not immutable. A growing body of empirical evidence from behavioral economics and cognitive neuroscience (e.g., Kahneman and Tversky (1979), Kahneman et al. (1991), Rabin (1998), Frederick et al. (2002), Camerer (2004), Camerer et al. (2005), Kahneman (2011), Gabaix (2019)) demonstrates that preferences can evolve systematically in response to experience, context, and internal cognitive and affective states. For instance, individuals may exhibit ambiguity aversion in unfamiliar situations (Ellsberg (1961)), display loss aversion when facing potential losses relative to a reference point (Kahneman and Tversky (1979)), or act in a time-inconsistent manner due to present bias (e.g., Laibson (1997), O'donoghue and Rabin (1999)). These findings suggest that preferences are not fixed inputs to the decision-making process but rather are themselves shaped by it. This evidence motivates the development of more granular and unified frameworks that can accommodate the endogenous evolution of preferences while maintaining analytical tractability.

This paper introduces **Neural Differential Utility (NDU)**, a novel framework for modeling the endogenous evolution of preferences. NDU departs from traditional models by positing that preferences are shaped by a dynamic interplay between cognitive and affective processes, subject to **resource constraints** (e.g., Gershman et al. (2015), Lieder and Griffiths (2020), Bhui et al. (2021)).[2] By "resource con-

[1] For instance, consider the observed shifts in risk aversion following market crashes (e.g., Malmendier and Nagel (2011)), or the changes in consumption patterns influenced by social trends (e.g., Abel (1990)). These examples, along with others such as the emergence of loss aversion or time-inconsistent behavior, highlight the limitations of models with fixed preferences.

[2] These resources can be thought of as analogous to measurable quantities such as time available for deliberation, cognitive load capacity (e.g., working memory limitations), and metabolic energy allocated

straints," we mean that individuals have limited cognitive and affective resources available for making decisions and regulating emotions, which influence how preferences are formed and updated (See Supplement for detailed discussions).

Specifically, we model preferences using deep residual networks (ResNets, see He et al. (2016)), a powerful class of deep neural networks (DNNs).[3] ResNets are characterized by their use of skip connections, which facilitate the training of very deep architectures (Srivastava et al. (2015)) and enable the approximation of highly complex, nonlinear functions (Li et al. (2022)).[4] One of the key benefits of ResNets lies in their recursive mathematical structure, which can be rigorously analyzed using controlled dynamic systems (e.g., Weinan (2017), Haber and Ruthotto (2017), Li et al. (2018), Weinan et al. (2019), Li et al. (2022)). In the context of NDU, ResNets provide a flexible and tractable way to represent the dynamic evolution of various preferences, capturing the non-linear interactions between consumption, preference state, and future utility. They are grounded in a two-dimensional framework of neural functioning (see Table I, also, e.g., Camerer et al. (2005), Kahneman (2011)), distinguishing between controlled/automatic and cognitive/affective processes. We provide a rigorous proof of this insight in Qi (2025d) by establishing a Universal Approximation Theorem for NDU within L^p spaces.[5]

---

to mental processes. The distinction between cognitive and affective resources is crucial. Cognitive resources are employed in deliberate, analytical thinking, while affective resources are involved in processing emotions and regulating impulses. These resources are likely not perfect substitutes and may interact in complex ways. For instance, a high cognitive load might deplete affective resources, making an individual more prone to impulsive choices. We explore this interaction further in Supplement.

[3] Deep neural networks (DNNs), see LeCun et al. (2015), are a powerful class of representation learning models inspired by the structure and function of the human brain. They consist of interconnected neurons organized in layers. These layers form a hierarchical representation (e.g., Brocas and Carrillo (2008), Gabaix and Graeber (2024)) of observations, with each layer learning increasingly abstract features. See Section 4 for comprehensive discussions of DNNs and ResNets.

[4] ResNets offer several advantages over other function approximators in this context. Their recursive structure naturally captures the dynamic nature of preference evolution, where past experiences and future expectations influence current choices. The use of skip connections allows ResNets to learn both gradual and abrupt changes in preferences, making them well-suited for modeling phenomena like shifts in risk aversion after market crashes or the sudden emergence of new desires. Furthermore, the layered architecture of ResNets can be interpreted as representing different levels of cognitive and affective processing, providing a link to the neuroeconomic foundations of the model.

[5] For instance, Camerer et al. (2005) highlight that **most behavior results from the interaction of all four quadrants** in Table I. The Universal Approximation Theorem demonstrates that, under specific

TABLE I

NDU AND TWO DIMENSIONS OF NEURAL FUNCTIONING

	Cognitive	**Affective**
Controlled Serial Effortful	I. Deliberate preference formation (e.g., planning, reasoning) Parameters optimization (θ_t)	II. Emotional regulation (e.g., resisting impulses) Parameters optimization (θ_t)
Automatic Parallel Effortless	III. Heuristic judgments, biases (e.g., framing effects) Preference state (u_t) dynamics	IV. Immediate emotional responses (e.g., fear, excitement) Preference state (u_t) dynamics

Note: This table presents a two-dimensional framework for understanding the psychological factors shaping preferences, drawing on insights from (e.g., Camerer et al. (2005), Kahneman (2011)). **Controlled** processes are conscious, deliberate, and resource-consuming. **Automatic** processes are fast, intuitive, and operate largely outside of conscious awareness. **Cognitive** influences relate to beliefs and knowledge, **affective** influences relate to emotions and moods. Each quadrant contributes to preference formation and provides a conceptual foundation for the NDU. Controlled processes shape the ResNets parameters θ_t, while automatic processes drive the dynamics of the preference state u_t.

We model a **resource-constrained agent** who initially lacks knowledge of the functional form and parameters of her best preferences model, and optimally adjusts her preference parameters, represented by the weights and biases θ_t of the ResNets. These adjustments, reflecting controlled cognitive and affective processes, incur preference adjustment costs (e.g., Lucas Jr (1967), Sims (2003)) representing the expenditure of limited cognitive and affective resources. Remarkably, in our examples, this cost structure promotes gradual preference shifts, leading to **endogenous risk and ambiguity aversion**[6] (see Proposition 1 and Qi (2025c)) and:

- **Cognitive resources** refer to the mental capacities for deliberation, planning, reasoning, and problem-solving. These resources are engaged when an agent consciously evaluates and adjusts their preferences or makes decisions based on new information or experiences. This aligns with the concept of "mental effort" or "cognitive control" discussed in the psychology literature (e.g.,

conditions on the ResNets architecture (e.g., sufficient depth and width), the activation function (e.g., ReLU), and the parameter space (e.g., compact), NDU can approximate a wide range of utility functions in $L^p([0,T];\mathbb{R}^n)$, including those exhibiting features like risk aversion, ambiguity aversion, loss aversion, and time inconsistency. Some economic models simplify by assuming cognition is the main driver and affect plays a relatively minor, automatic role. This assumption suggests that there are only two key dimensions (quadrants I and IV) at play, rather than the full four. In Qi (2025a), we show the controlled affective process can be the crucial element in generating loss aversion.

[6]For simplicity, our analysis of endogenous risk and ambiguity aversion considered the joint impact of cognitive and affective resources, rather than examining their individual contributions separately.

Botvinick et al. (2001), Kahneman (2011)). We illustrate this with an example of **endogenous time inconsistency** (see Qi (2025b)).

- **Affective resources** relate to the capacity for emotional regulation and management. Affective resources are deployed when an agent attempts to control or modify their emotional responses to stimuli, or when emotional states influence decision-making. This concept is related to the psychological constructs of "emotional intelligence" (e.g., Salovey and Mayer (1990), Mayer et al. (2008)) and "emotion regulation" (e.g., Gross (1998)). We illustrate this with an example of **endogenous loss aversion** (see Qi (2025a)).

Automatic processes are captured by a dynamic preference state u_t, which evolves according to a stochastic differential equation (SDE, see Definition 3) that is determined by the ResNets parameterized by θ_t. This captures the idea that some aspects of preference formation occur automatically and effortlessly (quadrants III and IV in Table I), such as immediate emotional responses or heuristic judgments. Note that the agent continues to hold **rational expectations**,[7] conditional on her current preferences and the dynamics of her preference state. That is, the agent maximizes recursive utility function given their current, endogenously evolving preferences:

$$V_t = \sup_{\theta_s} \mathbb{E}_t \left[\int_t^T \left(f_{\theta_s}(c_s, u_s, V_s) - \lambda \mathcal{L}(u_s, \theta_s) \right) ds \right]. \tag{1}$$

A key distinction of NDU is its emphasis on resource constraints. We model cognitive resources as pertaining to mental capacities for thinking and problem-solving, and affective resources as relating to emotional capacity and regulation. Adjusting the preference parameters θ_t requires cognitive or affective resources, incurring preference adjustment costs. These costs, analogous to a quadratic penalty in deep learning,[8] limit the agent's ability to instantaneously optimize her utility.

[7] The agent is assumed to have rational expectations in the sense that they correctly understand the dynamics of their preference state, u_t, and the stochastic process governing consumption, c_t. However, due to the presence of preference adjustment costs, the agent cannot instantaneously adjust their preferences to the optimal level. Therefore, their expectations are rational given their current preference parameters, θ_t, but they may not be fully rational in the traditional sense that ignores resource constraints.

[8] See Goodfellow (2016), Chapter 7 for a detailed discussion on regularization, also Section 3.2.

The optimization drives the agent's "best approximation", $V^*(\cdot;\theta)$, to asymptotically approach a (unknown) optimal target utility function $V(\cdot)$.[9] In the absence of resource constraints, preference adjustment costs become irrelevant, and $V^*(\cdot;\theta)$ corresponds to the traditional recursive utility framework (e.g., Kreps and Porteus (1978), Epstein and Zin (1989), Duffie and Epstein (1992)), where preference state u_t is exogenously determined, thereby providing a theoretical benchmark.

Why ResNets? According to McFadden (2001), the "Achilles heel" of economics lies in its lack of connection to underlying psychological and neurological mechanisms. This highlights the need for a comprehensive framework that links observed behavior to cognitive and affective processes. Intuitively, adopting neural networks to model these neurological mechanisms seems like a natural approach. ResNets (refer to Definition 2), with their recursive neural structure and skip connections, provide a natural way to model the dynamic interactions between different cognitive and affective processes, as depicted in Table I. Skip connections allow ResNets to learn both gradual and abrupt changes in preferences, making them well-suited for capturing the complexities of preference evolution. Notably, their layered representation of recursive decision processes enables mathematically rigorous analysis through controlled dynamic systems. This is crucial for ensuring the well-posedness and interpretability of the NDU framework.

Our central contribution is the development of the NDU framework itself, which offers a potentially unified approach to understanding preference evolution. **A closely related but distinct concept is rational inattention** (e.g., Sims (2003), Skiadas (2003), Maćkowiak et al. (2023)). Table II highlights the key differences between NDU and rational inattention models. While both NDU and rational inattention acknowledge cognitive limitations, they focus on different aspects of decision-making. Rational inattention primarily deals with how agents optimally acquire and process information when faced with cognitive constraints, leading to incomplete or

[9] This $V(\cdot)$ can be interpreted as the utility function the agent would exhibit if they could freely adjust their preferences without any resource constraints. It represents a "target" utility function that the agent strives to approximate, but the presence of adjustment costs prevents them from fully reaching it. The disparity between the optimal target utility function $V(\cdot)$ and its best approximation $V^*(\cdot;\theta)$ highlights the effect of limited resources on the agent's decision-making process.

noisy information about the state of the world (Simon (1955)). Agents in rational inattention models make optimal choices given their limited information, but their underlying preferences are typically assumed to be fixed. In contrast, NDU focuses on how cognitive and affective resource constraints shape the evolution of preferences themselves. It assumes that agents have full information about the state of the world but face limitations in their ability to adjust their preferences. Thus, NDU models preference formation, while rational inattention models belief formation.

TABLE II

Comparison of NDU Framework and Rational Inattention

Feature	**NDU Framework**	**Rational Inattention**
Primary Focus	Preference Formation	Belief Formation
Theoretical Basis	FBSDE System	Information Theory
Frictions	Resource Constraints	Cognitive Constraints
Belief	Full Information	Incomplete Information
Learning	Deep Learning	Bayesian Learning

Note: The rational inattention theory differs from NDU primarily by applying principles from information theory and concentrating on how information acquisition costs, driven by cognitive constraints, shape belief formation (Bayesian learning, see Baley and Veldkamp (2023) for detalied discussions). By contrast, NDU originates from deep learning theory (representation learning) and focuses on preference formation, incorporating both cognitive and affective costs. See Remark 2 and section 3 for details.

To further isolate the primary mechanism of NDU from rational inattention, our framework intentionally omits discussions of belief distortion.[10] This separation establishes a distinct demarcation and lays the benchmark for subsequent research to explore the interactions between the belief formation and preference formation.

Comparison with Multi-Prior Preferences and Robustness. Table III compares NDU with multi-prior preferences and robustness, three recursive frameworks for decision-making under uncertainty. NDU models preference formation as a deep

[10]Existing literature has shown that belief distortion can influence and reshape individual preferences. For instance, the recursive preference has been extended to accommodate multiple-priors (e.g., Gilboa and Schmeidler (1989), Schmeidler (1989), Epstein and Wang (1994)) through the incorporation of belief distortion via g-expectations Chen and Epstein (2002) and G-expectations Epstein and Ji (2013), also robustness and variational preference (e.g., Hansen and Sargent (2001), Maccheroni et al. (2006), Strzalecki (2011)). Furthermore, recent work by Borovička (2020) has established a connection between recursive preferences and the literature on endogenous discounting, shedding light on the formation of heterogeneous beliefs, a research agenda pioneered by Koopmans (1960), Uzawa (1968).

learning process of "best approximation" using ResNets and a state variable, constrained by resources. Multi-Prior Preferences tackle ambiguity aversion by considering the worst-case scenario across a set of probability distributions. Robustness addresses model uncertainty by optimizing for the worst-case across a set of unstructured alternative models. While NDU uses deep learning to optimizing preferences, Multi-Prior Preferences involve implicit learning through prior updating. Robustness doesn't focus on learning but on finding decisions that perform robust across various unstructured models. Essentially, NDU learns the features of preferences, while the others deal with uncertainty about probabilities or models.[11]

TABLE III

COMPARISON OF NDU, MULTI-PRIOR PREFERENCES, AND ROBUSTNESS

Feature	**NDU**	**Multi-Prior Preferences**	**Robustness**
Focus	Preference Formation	Ambiguity Aversion	Model Uncertainty
Unknown	Preferences Model	Probability Distributions	Unstructured Alternatives
Mechanism	**Best Approximation**	Worst-Case	Min-Max Optimization
Key Elements	$f_{\theta_t}(c_t, u_t, V_t)$	Set of Priors $\mathcal{P}$	Set of Models $\mathcal{Q}$
Learning	Deep Learning	Implicit (Prior Updating)	-
Framework	FBSDE, ResNets	BSDE/Maxmin EU	Robust Control
Belief Distortion	-	Multiple Priors	Variational Preferences
Literature	This paper	Chen and Epstein (2002)	Hansen and Sargent (2001)

Note: NDU focuses on how resource constraints shape the evolution of preferences represented by ResNets parameters θ_t and state u_t. Multi-prior preferences (e.g., Gilboa and Schmeidler (1989), Schmeidler (1989), Epstein and Wang (1994), Chen and Epstein (2002), Epstein and Ji (2013)) address ambiguity aversion by considering a set of priors $\mathcal{P}$. Robustness emphasizes decisions that are robust to model misspecification (e.g., Hansen and Sargent (2001), Skiadas (2003), Maccheroni et al. (2006), Strzalecki (2011)) represented by a set of alternative unstructured models $\mathcal{Q}$. While distinct, these frameworks can be complementary. For the definition of g-expectation, refer to Peng (1997), and G-expectation, see Peng (2007). EU denotes Expected Utility.

1.1. *The NDU Framework*

Consider an agent with a finite time horizon $T > 0$ who chooses preference parameter controls θ_s. The agent's utility at time s, denoted by V_s, is governed by the

[11] These frameworks can be seen as complementary to each others. For instance, one could incorporate multi-prior or robustness considerations into the NDU framework by allowing the agent to be uncertain about the parameters of their own preference evolution process or by considering a set of possible ResNets specifications. This would lead to a richer model that captures both the dynamics of preference formation and the agent's uncertainty about those dynamics.

following controlled forward-backward stochastic differential equation (FBSDE)[12] system (see Section 3.1 for details):

$$du_t = h_{\theta_t}(u_t, c_t)dt + \sigma^h_{\theta_t}(u_t, c_t)d\mathcal{Z}^h_t, \quad u_0 = \phi(c_0), \qquad \text{(Forward Propagation)}$$

$$dV_s = -[f_{\theta_s}(c_s, u_s, V_s) - \lambda\mathcal{L}(u_s, \theta_s)]ds + \sigma_s d\mathcal{Z}_s, \quad V_T = 0, \qquad \text{(Backward Propagation)}$$

where we discuss the key insights of this specific FBSDE system below. The forward equation (Forward Propagation)[13] describes the evolution of the preference state u_t, driven by the ResNets function h_{θ_t}, a stochastic term $\sigma^h_{\theta_t}(u_t, c_t)d\mathcal{Z}^h_t$, and the initial preference mapping $\phi(c_0)$. The backward equation (Backward Propagation) describes the evolution of the value function V_s, which is determined by the aggregator function f_{θ_s}, the preference adjustment costs function $\mathcal{L}$, and a stochastic term $\sigma_s d\mathcal{Z}_s$. This system captures the dynamic interplay between the agent's current consumption, her evolving preference state, and her expectations about future utility. The agent's control problem is to choose the optimal parameter trajectory θ to maximize her expected utility, net of preference adjustment costs.

Key Insights. The NDU framework models the evolution of preferences over time as agents balance immediate desires (local optimality) with long-term goals (global optimality), analogous to training a deep learning model (see Section 4.4 for details). This involves adjusting "preference parameters" θ_t that govern the valuation of consumption and future utility, subject to cognitive and affective resource constraints. The framework posits a dynamic interaction between forward and backward propagation. To best illustrate the basic intuition of the NDU framework, we adopt the following simple case to highlight the key insights:

---

[12] Classical forward-backward stochastic differential equation (FBSDE) is examined in Peng and Wu (1999), and this form of FBSDE emerges in various mean-field type problems, including mean-field games (MFG) and mean-field type control problems, as demonstrated in Carmona and Delarue (2015) and Ahuja (2016). The concept notably shares a deep and relevant connection with the work on deep learning theory presented in Weinan et al. (2019). Additionally, the Forward Propagation also has connections to recursive utility with habit formation, which is discussed in Section 2.2 and Remark 3.

[13] The forward equation is also known as the Fokker-Planck (Kolmogorov forward) equation in MFG.

1. **Forward Propagation: Forming Current Preference State u_t.** We consider a simple scenario where the agent's preference state u_t is a linear function of current consumption, modulated by the preference parameter θ_t:

$$u_t = \theta_t c_t. \qquad \text{(Linear Forward Propagation)}$$

Here, θ_t represents the intensity with which current consumption at time t affects the preference state. This can be interpreted as an automatic, low-deliberation cognitive and affective process (quadrants III and IV in Table I), where θ_t acts as a weight determining the salience of consumption in shaping preferences. This is analogous to a **single-layer linear neural network**,[14] where θ_t corresponds to the network weight. See Remark 3 for detailed discussion.

2. **Backward Propagation: Preference Adjustment and Expected Utility.** The agent adjusts θ_s to maximize expected future utility (quadrants I and II in Table I). For simplicity, we consider an infinite horizon and assume stationarity. We further assume quadratic preference adjustment costs, dependent on both θ_s and c_s: $\mathcal{L}(u_s, \theta_s) = \frac{1}{2} c_s \theta_s^2$. The backward propagation equation becomes:

$$dV_s = -\left[\theta_s c_s - \beta V(c_s) - \frac{\lambda c_s \theta_s^2}{2}\right] ds + \sigma_s d\mathcal{Z}_s, \qquad \text{(Linear Backward Propagation)}$$

where $\beta > 0$ is the discount rate and $\lambda > 0$ scales the preference adjustment costs. The term $\theta_s c_s$ represents the instantaneous utility derived from consumption at time s, influenced by the current preference parameter. The preference adjustment costs $\frac{\lambda c_s \theta_s^2}{2}$ reflect the cognitive or affective resources expended in modifying preferences at time s, which now depend on the current consump-

[14] In this simplified example (see Remark 3), the preference state u_t can be thought of as a scalar representing the agent's overall level of satisfaction or utility derived from consumption. The parameter θ_t determines how much weight the agent places on current consumption when forming her preference state. A higher value of θ_t implies that current consumption has a stronger influence on the agent's preferences. We define the ResNets (see Definition 3) in the context of dynamic preference states.

tion level, introducing a novel state-dependent cost.[15] This adjustment process is analogous to the backward pass in a neural network, where the agent (or the "network") tunes the weights θ_s to improve the outcome (expected utility), but this tuning incurs costs that are dependent on the consumption level.

3. **Aggregator Function and Integrated Influences** The aggregator function $f_{\theta_s}(c_s, u_s, V_s)$ in the backward propagation integrates consumption c_s, preference state u_s, and future utility V_s into instantaneous utility at time s, similar to a neural network's output layer. In this simplified case, the aggregator becomes $f_{\theta_s}(c_s, u_s, V_s) = \theta_s c_s - \beta V(c_s) - \frac{\lambda c_s \theta_s^2}{2}$. This function, parameterized and adjusted by θ_s, reflects the combined influence of automatic preference formation (through $u_s = \theta_s c_s$) and controlled, effortful adjustment of preferences (through the cost $\frac{\lambda c_s \theta_s^2}{2}$). It shows how agents dynamically value consumption and future utility within resource constraints, with preference adjustment costs now intertwined with the current consumption level.

This linear example illustrates the core mechanism of NDU: the dynamic interaction between forward propagation (preference state formation) and backward propagation (preference adjustment) under resource constraints. The general NDU framework extends this by allowing for non-linear relationships through the ResNets function h_{θ_t} and the aggregator function f_{θ_s} (see Section 3).

Proposition 1—Endogenous Risk Aversion: *Consider the NDU framework as defined by equations* (Linear Forward Propagation) *and* (Linear Backward Propagation). *An agent maximizes expected lifetime utility by dynamically choosing the preference parameter sequence* $\{\theta_s\}_{s\in[t,\infty)}$ *from time t onward. The consumption (data generating) process,* c_t*, follows the geometric Brownian motion:*

$$dc_t = \mu c_t dt + \sigma c_t d\mathcal{Z}_t, \quad c_0 > 0, \tag{DGP}$$

[15] This captures the idea that changing preferences is more costly when current consumption is high, perhaps because it requires more cognitive effort to override strong habits or immediate desires. The backward equation describes how the agent's expected utility evolves over time, taking into account the instantaneous utility from consumption, the discounted future utility, and the cost of adjusting preferences. The agent's goal is to choose the preference parameter θ_s to maximize her expected utility, net of the adjustment costs.

where μ denotes the growth rate, $\sigma > 0$ is the volatility, and $\mathcal{Z}_t$ is a standard Brownian motion. The agent's objective is to find the optimal policy $\{\theta_s^*\}_{s\in[t,\infty)}$ to maximize the objective function starting from time t:

$$\sup_{\{\theta_s\}_{s\in[t,\infty)}} \mathbb{E}_t\left[\int_t^\infty e^{-\beta(s-t)}\left(\theta_s c_s - \frac{1}{2}\lambda c_s \theta_s^2\right) ds\right], \tag{2}$$

where $\beta > 0$ is the discount rate, $\lambda > 0$ is the preference adjustment costs parameter. Assume the value function $V(c)$ satisfies $V(1) = 0$ and the transversality condition $\lim_{t\to\infty} \mathbb{E}_t[e^{-\beta(s-t)}V(c_s)] = 0$. The Hamilton-Jacobi-Bellman (HJB) equation is given by

$$0 = \sup_\theta \left\{\theta c - \frac{1}{2}\lambda c\theta^2 - \beta V(c) + \mu c V'(c) + \frac{1}{2}\sigma^2 c^2 V''(c)\right\}. \tag{3}$$

Then, under the condition $\beta > \mu$, the optimal preference parameter is given by $\theta^*(c) = \frac{1}{\lambda}$, and the value function is:

$$V(c) = \frac{c - c^\gamma}{2\lambda(\beta-\mu)}, \quad \textit{with} \quad \gamma = \frac{-(\mu - \frac{1}{2}\sigma^2) - \sqrt{(\mu - \frac{1}{2}\sigma^2)^2 + 2\sigma^2\beta}}{\sigma^2} < 0. \tag{4}$$

The coefficient of relative risk aversion, $\mathcal{R}(c) = -\frac{cV''(c)}{V'(c)}$, approaches $1-\gamma > 1$ as $c \to 0$, indicating significant risk aversion, and $\mathcal{R}(c) \to 0$ as $c \to \infty$. To better understand the value function in equation (4), we define $u(c) = V'(c)$, where $u(c) = \frac{1-\gamma c^{\gamma-1}}{2\lambda(\beta-\mu)}$. Analyzing $u(c)$'s derivatives, we obtain a local measure of relative risk aversion: $\rho = -c \cdot \frac{u''(c)}{u'(c)} = 2 - \gamma$. This suggests a linkage to Constant Relative Risk Aversion (CRRA) utility.

Remark 1—Implications of Proposition 1: This proposition demonstrates that even with a simple linear ResNets and quadratic preference adjustment costs, NDU can generate endogenous risk aversion.[16] The optimal preference parameter, $\theta^*(c) = \frac{1}{\lambda}$, is inversely related to the preference adjustment costs, λ. Higher

[16] Another analogous interesting parallel in corporate finance is found by Bolton et al. (2011), which shows that a financially constrained firm, even if initially risk-neutral, may exhibit risk-averse behavior and engage in endogenous risk management strategies when it lacks adequate cash reserves.

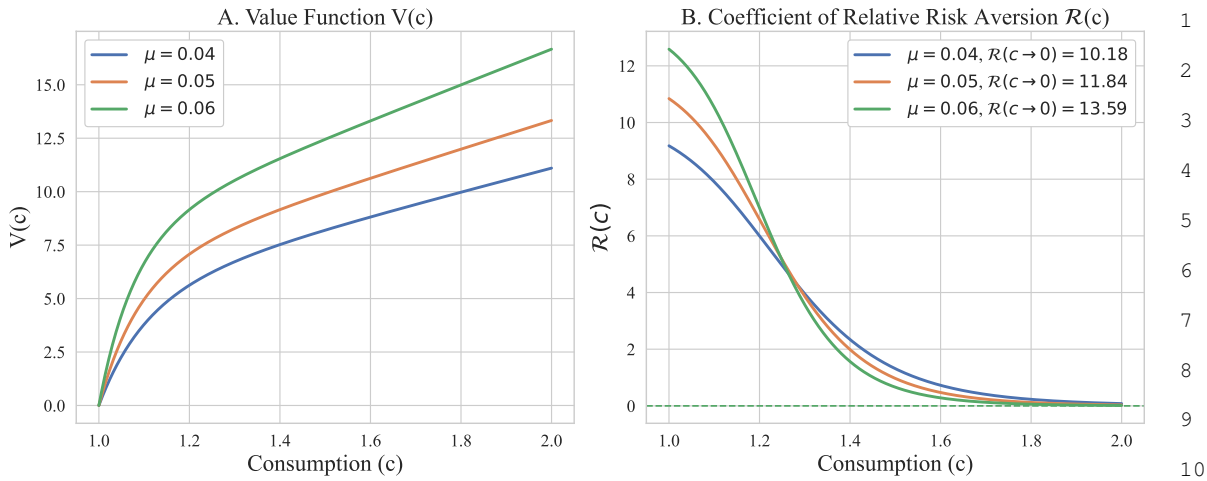


FIGURE 1.—**Endogenous Risk Aversion.** Panel A displays the value function $V(c)$, and Panel B displays the coefficient of relative risk aversion $\mathcal{R}(c)$, both plotted against consumption c. The fixed parameters used are $\sigma = 0.1$, $\beta = 0.1$, and $\lambda = 1.5$. The dashed lines in the Panel B indicate the limits of $\mathcal{R}(c)$ as c approaches ∞. As consumption approaches zero, the agent becomes highly risk-averse, with the coefficient of relative risk aversion approaching $1 - \gamma > 1$. As consumption increases, the agent becomes less risk-averse, with the coefficient approaching 0. This pattern is consistent with empirical evidence suggesting that individuals are more risk-averse at low levels of wealth or consumption.

preference adjustment costs make the agent less willing to change her preferences in response to consumption fluctuations, effectively leading to risk-averse behavior. The proposition also reveals decreasing relative risk aversion (DRRA), with $\mathcal{R}(c)$ approaching $1 - \gamma > 1$ as consumption approaches zero and 0 as consumption goes to infinity. This means that agents are more risk-averse at low consumption levels, consistent with empirical evidence suggesting that wealthier individuals tend to be more risk-tolerant (e.g., Rabin (2000), Ogaki and Zhang (2001), Post et al. (2008), Swanson (2012)). Unlike traditional models where risk aversion is imposed through concave utility, NDU endogenizes it as an outcome of optimal preference adjustment under uncertainty and costly adjustment. This suggests model predictions: individuals facing higher costs of preference adjustment may exhibit greater risk aversion, particularly at low consumption levels. The NDU framework offers a new perspective on risk aversion, linking it to the dynamics of preference formation.

Summary. The NDU framework utilizes the forward and backward propagation structure inherent in ResNets to model the evolution of preferences. Forward

propagation characterizes the automatic evolution of the preference state, while backward propagation captures the controlled, resource-constrained optimization and preference parameters adjustment. This approach generalizes the recursive utility framework of (e.g., Kreps and Porteus (1978), Epstein and Zin (1989), Duffie and Epstein (1992)) in three key ways:

1. **Dynamic Aggregator Function:** NDU replaces the traditional, pre-specified aggregator $f(c, v)$ with a dynamic aggregator $f_{\theta_s}(c, u, v)$ controlled by a set of parameters θ_s. This allows the functional form of the aggregator, and thus the agent's preferences, to evolve over time. The inclusion of the preference state u further expands the model's expressiveness by incorporating contextual factors influencing utility.
2. **Endogenous Preference Formation:** NDU models the evolution of preferences via the FBSDE system with controlled parameters θ_s. This stands in contrast to traditional recursive utility where preferences, embodied in the fixed aggregator function. The optimization process of θ_s enables past experiences and anticipated future outcomes to influence current preferences, thereby allowing NDU to capture the impact of contextual factors on decision-making processes.
3. **Universal Approximation Theorem:** Building upon the *Universal Approximation Theorem*[17] for standard ResNets, we establish conditions for NDU to endogenize various well-defined utility functions within the L^p spaces (see Qi (2025d)). This approach provides a more generalized framework compared to standard recursive utility models, as it encompasses the later as a special case, assuming θ_s remains constant and u_t is exogenously determined as $u(c)$.

The remainder of this paper is organized as follows: Section 2 provides an introduction to ResNets. Section 3 defines the NDU. Section 4 outlines the properties

[17] Li et al. (2022) demonstrate that ResNets can approximate any continuous function on a compact domain arbitrarily well, given sufficient depth and width. They idealize ResNets as continuous-time dynamical systems governed by ordinary differential equations (ODEs) and establish sufficient conditions on the control families driving these ODEs that ensure universal approximation in L^p spaces. This framework relies on "flow maps," generated by ODEs driven by appropriate control families, to represent non-linear relationships between inputs and outputs. In contrast, this paper generalizes the standard ResNets structure, which is based on ODEs, to incorporate stochastic consumption processes, modeled using SDEs. This approach is connected to Deep Q-Networks (DQN, see Mnih et al. (2015)), Qi (2025d) also generalize the Universal Approximation of standard ResNets to DQN in the NDU context.

and axioms underlying the utility. Section 5 concludes with AI implications. Main proofs are in Appendix. Additional information can be found in the Supplement.

2. PRELIMINARIES

This section establishes the foundation for NDU. We review recursive utility, particularly Stochastic Differential Utility (SDU, Duffie and Epstein (1992)), and then formulate a stochastic dynamical systems approach to generalize the standard Deep Residual Networks (ResNets, He et al. (2016), Weinan (2017), Haber and Ruthotto (2017), Li et al. (2018), Weinan et al. (2019), Li et al. (2022)) to incorporate the stochastic consumption process, highlighting its connection to mean-field optimal control. We define continuous-time ResNets as a limit of discrete-time ResNets.

2.1. *Notations*

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions: the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is right-continuous and $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets, with $T > 0$ as the finite time horizon. An agent chooses a consumption path $(c_t)_{t\in[0,T]}$, where $c_t \in \mathcal{C} \subset \mathbb{R}^n$ is the consumption bundle at time t, measured in units of a numeraire good. The agent's utility is represented by a utility function $V : \mathcal{C} \times [0,T] \to \mathbb{R}$, which we aim to approximate using deep learning. We consider a specific class of ResNets architectures with L layers and element-wise activation function $\eta(\cdot)$. Let $\Theta \subset \mathbb{R}^d$ be a compact parameter space. The hypothesis space $\mathcal{H}$ consists of functions parameterized by these ResNets:

$$\mathcal{H} = \left\{ V(\cdot;\theta) : \mathcal{C} \to \mathbb{R} \mid \theta \in L^\infty([0,T];\Theta) \right\}, \tag{5}$$

where $\theta := \{\theta_t\}_{t\in[0,T]}$ represents the trajectory of network parameters, and θ_t denotes the parameter vector at time t. $L^\infty([0,T];\Theta)$ denotes the space of essentially bounded parameter trajectories. Given a probability measure P on $\mathcal{C}$ representing the consumption distribution, the optimal approximant $V^* \in \mathcal{H}$ minimizes the expected loss:

$$V^*(c;\theta) = \inf_{V(\cdot;\theta)\in\mathcal{H}} \int_{\mathcal{C}} \ell(V(c), V(c;\theta)) dP(c), \tag{6}$$

TABLE IV
NOTATION SUMMARY

Notation	Description	Domain/Range
c_t	Consumption bundle at time t	$\mathcal{C} \subset \mathbb{R}^n$
u_t	Preference state at time t	$\mathbb{R}^q$
θ_t	ResNets parameters at time t	$\Theta \subset \mathbb{R}^d$
V_t^P	Stochastic Differential Utility (SDU) at time t	$\mathbb{R}$
$V_t(c;\theta)$	NDU at time t given c and θ	$\mathbb{R}$
$f(c,h,V)$	Recursive utility aggregator with habit formation	$\mathcal{C} \times \mathcal{C} \times \mathbb{R} \to \mathbb{R}$
$f_{\theta_s}(c_s,u_s,V_s)$	NDU aggregator function	$\mathcal{C} \times \mathbb{R}^q \times \mathbb{R} \to \mathbb{R}$
$h_{\theta_t}(u_t,c_t)$	Residual block function	$\mathbb{R}^q \times \mathcal{C} \to \mathbb{R}^q$
$\eta(\cdot)$	Activation function	$\mathbb{R}^m \to \mathbb{R}^m$
$\phi(\cdot)$	Initial preference mapping	$\mathcal{C} \to \mathbb{R}^q$
$\ell(V,\widehat{V})$	Loss function	$\mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$
$\mathcal{L}(u,\theta)$	Preference adjustment cost	$\mathbb{R}^q \times \Theta \to \mathbb{R}$
$\mathbb{L}$	Number of layers in ResNets	$\mathbb{N}$

where $\ell : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ is a continuous and convex loss function, and $V(c)$ represents the (unknown) optimal target utility derived from consumption path c.[18]

REMARK 2—Comparison with Robustness and Rational Inattention: NDU, variational preferences (Hansen and Sargent (2001), Skiadas (2003), Maccheroni et al. (2006), Strzalecki (2011)), and rational inattention models (Sims (2003)) all address decision-making under uncertainty and evolving preferences, but they differ fundamentally in their methodologies and focuses. Variational preferences address ambiguity aversion by minimizing expected utility plus a penalty for deviating from a reference probability measure:

$$\inf_{Q \in \mathcal{Q}} \left\{ \int_\Omega \mathbb{E}_Q[u(c)] + \kappa D_{KL}(Q||P) \right\}, \qquad \text{(Robustness)}$$

[18] The existence of a minimizer $V^* \in \mathcal{H}$ is guaranteed under the compactness of Θ, continuity of ℓ, and uniform bounds on network weights (see, e.g., Anthony et al. (1999)). These conditions, along with the continuity of V in θ, ensure the closedness of the hypothesis space and the continuity of the objective function in (6). However, uniqueness may require additional assumptions, such as strict convexity of the mapping $\theta \mapsto V(\cdot,\cdot;\theta)$ or a regularization term that ensures strict convexity. We will introduce such a term later in the form of a preference adjustment cost. In practice, even if uniqueness is not guaranteed, optimization algorithms can often find good local minima that provide useful approximations.

where $\mathcal{Q}$ is a set of alternative models, u is an instantaneous utility function, $\kappa > 0$ is a penalty parameter, and $D_{KL}(\cdot||\cdot)$ is a divergence measure (e.g., Kullback-Leibler divergence). This emphasizes the agent's search for a robust belief in the face of model uncertainty. Rational inattention models focus on constraints on information processing capacity. Agents optimally choose the precision of signals about the state of the world, subject to an information cost typically related to entropy reduction:

$$\max_{S} \{\mathbb{E}[u(c|S)] - \lambda_I I(S;X)\}, \qquad \text{(Rational Inattention)}$$

where S is the signal, X is the state variable, $I(S;X)$ is the mutual information between the signal and the state, $\pi(S|X)$ represents the information structure and $\lambda_I > 0$ represents the cost of information. The focus is on the trade-off between the benefits of more precise information and the cost of acquiring it. The NDU framework, in contrast, aims to directly learn the evolution of preferences from observed consumption choices by minimizing a structural risk functional. This functional balances the empirical fit to the data with a regularization term that controls the complexity of the learned preference dynamics. Let $\{c_i\}_{i=1}^N$ be a set of observed consumption paths. The NDU framework minimizes:

$$\min_{\theta \in L^\infty([0,T];\Theta)} \left\{ \frac{1}{N} \sum_{i=1}^{N} \ell(V(c_i), V(c_i;\theta)) + \lambda \int_0^T \mathcal{L}(u_s, \theta_s) ds \right\}, \qquad \text{(NDU)}$$

where $V(c_i;\theta)$ approximates the (unknown) optimal target $V(c_i)$ at time 0 using a neural networks parameterized by $\theta \in L^\infty([0,T];\Theta)$, $\ell(\cdot,\cdot)$ is a loss function, $\lambda \geq 0$ is a regularization parameter, and $\mathcal{L}(\cdot;\theta)$ is a function that penalizes the complexity of the network weights evolving over time, interpreted as a preference adjustment cost. The optimization occurs within the space of bounded functions $L^\infty([0,T];\Theta)$. NDU directly learns preference dynamics from data, while the regularization term controls model complexity, analogous to controlling for model risk or information processing costs in the other approaches (see Remark 5 for discussions).[19]

[19] NDU can be related to variational preferences and rational inattention in specific cases. For instance, if the preference adjustment cost $\mathcal{L}(u,\theta)$ in NDU is linked to the divergence between probability

2.2. *Recursive Utility Paradigm*

We consider a discrete-time setting with periods $t = 0, 1, \ldots, T$ and consumption $c_t \in \mathcal{C} \subset \mathbb{R}^n$. A generalization of the standard intertemporally additive expected utility model is the recursive utility specification (e.g., Kreps and Porteus (1978), Epstein and Zin (1989)):

$$V_t = f(c_t, \mathbb{E}_P[V_{t+1}|\mathcal{F}_t]), \tag{7}$$

where the aggregator $f : C \times \mathbb{R} \to \mathbb{R}$ is strictly increasing and concave in its first argument and strictly increasing in its second argument. This can incorporate habit formation (e.g., Duffie and Epstein (1992), Duffie and Skiadas (1994)) by making the aggregator depend on a habit stock h_t, which typically evolves as a function of past consumption, e.g., $h_t = (1-\alpha)h_{t-1} + \alpha c_{t-1}$, where $\alpha \in (0,1)$. Then we can write the aggregator as $f(c_t, h_t, \mathbb{E}_t[V_{t+1}])$. Here, $\mathbb{E}_t[\cdot]$ represents the conditional expectation operator given the information set $\mathcal{F}_t$ at time t, under the probability measure P. We now define SDU in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ satisfying the usual conditions, where $T > 0$ is a finite time horizon.[20]

Definition 1—Stochastic Differential Utility with Habit Formation: For $c, h \in \mathcal{D}$, the SDU process $(V_t^P)_{t\in[0,T]}$ is the unique solution to the following BSDE:

$$dV_t^P = -f(c_t, h_t, V_t^P)dt + \sigma_t d\mathcal{Z}_t^P, \quad V_T^P = 0, \tag{8}$$

where:

- $f : \mathcal{C} \times \mathcal{C} \times \mathbb{R} \to \mathbb{R}$ is the aggregator function, assumed to be Lipschitz continuous, and increasing and concave in its first argument.
- $(\mathcal{Z}^P)_{t\in[0,T]}$ is a d-dimensional general martingale on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$, representing the source of uncertainty (Duffie and Epstein (1992)).

---

measures or information processing costs, we might recover structures similar to cases Robustness or Rational Inattention. However, such connections are not the primary focus of NDU in the current paper.

[20] Stochastic Differential Utility (SDU) can be motivated as a continuous-time approximation of this discrete-time model (see Kraft and Seifried (2014)), complementing its original definition by Duffie and Epstein (1992). We allow for a general filtration $(\mathcal{F}_t)_{t\in[0,T]}$, enabling us to model a wide range of uncertainty structures beyond the typical Brownian filtration.

- $\sigma : [0,T] \times \Omega \to \mathbb{R}^{1\times d}$ is a progressively measurable process satisfying $\mathbb{E}[\int_0^T \|\sigma_t\|^2 dt] < \infty$, where $\|\cdot\|$ denotes the Frobenius norm.

The SDU process V_t^P represents the agent's utility at time t for the processes c and h, conditional on $\mathcal{F}_t$. While SDU with habit formation is more realistic than additive expected utility, specifying a suitable aggregator f remains challenging, motivating our use of flexible ResNets to approximate it within distinct scenarios.

2.3. *ResNets and Mean-field Optimal Control*

We consider a hypothesis space $\mathcal{H}$ consisting of neural network functions with varying architectures. To illustrate the economic mechanism of deep learning in representing preferences, we initially focus on approximating the instantaneous utility with a recursive structure. We begin with the standard aggregator $f(c,v) = u(c) - \beta v$ as a starting point, where $u(c)$ is the instantaneous utility from consumption c and $\beta \in (0,1)$ is a discount rate.[21]

Definition 2—Deep Residual Networks: Consider a discrete time grid $0 = t_0 < t_1 < ... < t_{\mathbb{L}} = T$ with step size $\Delta t = T/\mathbb{L}$. A feed-forward deep residual network with $\mathbb{L}$ layers is a function $\hat{u}_{\mathbb{L}} : \mathcal{C} \to \mathbb{R}^q$ parameterized by $\theta = (\theta_0, \ldots, \theta_{\mathbb{L}-1})$, where:

- $\mathcal{C} \subset \mathbb{R}^n$ represents the input space (consumption bundles).
- $\mathbb{R}^q$ represents the output space (preference state), where q is the number of latent factors influencing preferences.
- Each $\theta_l \in \Theta$ represents the parameters of the l-th layer, where $\Theta \subset \mathbb{R}^d$ is the parameter space. The function h_{θ_l} can be a multi-layer neural network itself with parameters θ_l.

The output of the network is given recursively by:

$$u_{l+1} = u_l + \Delta t \cdot h_{\theta_l}(u_l, c_l), \quad l = 0, 1, \ldots, \mathbb{L}-1, \tag{9}$$

[21] ResNets can effectively represent dynamic preferences due to their ability to capture complex, non-linear transformations of a preference state over time. Each layer of the ResNets can be interpreted as a stage in the evolution of preferences, where the "hidden" nodes represent latent factors influencing preferences. These factors could include psychological biases, cognitive limitations, or evolving tastes. The weights connecting the nodes represent the strength of influence between these factors, and the activation function introduces non-linearities, allowing for complex interactions between factors.

where $u_0 = \phi(c_0)$ is the initial preference state, and each function $h_{\theta_l} : \mathbb{R}^q \times \mathcal{C} \to \mathbb{R}^q$ specifies the architecture of the network, called a residual block, and is parameterized by a feedforward neural network with parameters θ_l at each layer l.

Now, we take the limit as $\Delta t \to 0$ to obtain the continuous-time ResNets.

Definition 3—Continuous-time ResNets: A continuous-time ResNets is a stochastic process u_t taking values in $\mathbb{R}^q$, driven by an adapted stochastic process $c_t \in \mathcal{C} \subset \mathbb{R}^n$, and parameterized by a measurable, right-continuous, and uniformly bounded parameter function $\theta_t \in \Theta \subset \mathbb{R}^d$. The evolution of u_t is governed by a controlled stochastic differential equation (SDE):

$$du_t = h_{\theta_t}(u_t, c_t)dt + \sigma^h_{\theta_t}(u_t, c_t)d\mathcal{Z}^h_t, \quad u_0 = \phi(c_0), \tag{10}$$

where:

- $h_{\theta_t} : \mathbb{R}^q \times \mathcal{C} \to \mathbb{R}^q$ is a function that specifies the architecture of the network. It is parameterized by θ_t. Economically, h_{θ_t} governs the deterministic evolution of the preference state, capturing how preferences change based on current consumption and the existing preference state.
- $\sigma^h_{\theta_t} : \mathbb{R}^q \times \mathcal{C} \to \mathbb{R}^{q\times m}$ is a progressively measurable and uniformly bounded volatility matrix, with dimension m depending on the dimension of the martingale $\mathcal{Z}^h$. This term captures random shocks on preference evolution.
- $\mathcal{Z}^h_t$ is an m-dimensional martingale process on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$, representing the source of randomness in preference dynamics.
- $\phi : \mathcal{C} \to \mathbb{R}^q$ determines the initial value of the process u_t based on the initial consumption c_0. This function maps the initial consumption experience to an initial preference state.

This continuous-time SDE for the preference state can be approximated by a deep discrete-time ResNets with many layers and small time steps, as shown in the following proposition. This is a standard result in the numerical analysis of SDEs.

Proposition 2: *Suppose Assumptions 5 and 6 hold. Then, for a given parameter trajectory $\theta \in L^\infty([0,T];\Theta)$, the solution to the SDE (10) is unique and can be characterized by*

a flow map:

$$\Phi_T(\cdot;\theta) : \mathcal{C} \to \mathbb{R}^q, \tag{11}$$

defined as $\Phi_T(c_0;\theta) = u_T(c_0,\theta)$*, where* u_t *is the solution to* (10) *at time* t *with initial condition* c_0 *and parameter trajectory* θ*. This map depends on the initial consumption bundle* c_0 *and the parameter trajectory* θ*, capturing the learned transformation of the ResNets over time. If* ϕ *is continuous, and* h_{θ_t} *is Lipschitz continuous in* (u,c) *uniformly in* t *and* θ*, then the flow map* Φ_t *is jointly continuous in* c_0 *and* θ *with respect to the Euclidean norm on* $\mathcal{C}$ *and the* L^∞ *norm on* $L^\infty([0,T];\Theta)$*, for each fixed* t*.*

The continuous-time ResNets, represented by the flow map Φ_T, models the evolution of an agent's preference state over time. Each infinitesimal transformation of the preference state is defined by the residual block:

$$h_{\theta_t}(u_t, c_t) = A_t\eta(B_t u_t + \Gamma_t c_t + b_t), \tag{12}$$

where $\theta_t = (A_t, B_t, \Gamma_t, b_t)$ with $A_t \in \mathbb{R}^{q\times m}$, $B_t \in \mathbb{R}^{m\times q}$, $b_t \in \mathbb{R}^m$, and $\eta : \mathbb{R}^m \to \mathbb{R}^m$ is an element-wise activation function. This block can be interpreted as a continuous-time transformation of the preference state through m neurons. The matrix $\Gamma_t \in \mathbb{R}^{m\times n}$ introduces a direct dependence of the preference state dynamics on the consumption path. See Section 4.3 for discussions of the activation function.

Mean-field Optimal Control. The parameters θ_t in the NDU model are optimized by formulating a mean-field optimal control problem (see Section 5).[22] We seek a parameter trajectory θ that maximizes the utility derived from a consumption path while regularizing the complexity of the parameter dynamics. Consider a probability space $(\Omega, \mathcal{F}, P)$ and a consumption path $c_s : \Omega \to \mathcal{C}$ for each $s \in [0,T]$, where c_s is a random variable adapted to a filtration $\{\mathcal{F}_s\}_{s\in[0,T]}$. The mean-field

[22] This approach (e.g., Li et al. (2018), Weinan et al. (2019)) is suitable for modeling preference evolution because it allows us to consider the aggregate behavior of a large population of interacting "neurons" or cognitive processes, represented by the parameters θ_t. By assuming that parameters are chosen optimally ex-ante, we are essentially capturing the idea that individuals have a long-run preference structure that they strive to maintain, subject to adjustment costs (See Section 3.2 for discussions).

control problem can be formulated as:

$$V_t(c;\theta) = \sup_{\theta_s \in L^\infty([0,T];\Theta)} \mathbb{E}\left[\int_t^T e^{-\beta(s-t)} \left(\Phi_s(c_s;\theta) - \lambda \mathcal{L}(u_s,\theta_s)\right) ds | \mathcal{F}_t\right], \tag{13}$$

subject to (10), where:

- $\mathbb{E}[\cdot \mid \mathcal{F}_t]$ denotes the conditional expectation operator.
- $\Phi_s(c_s;\theta) : \mathcal{C} \to \mathbb{R}$ represents the instantaneous utility at time s derived from consumption c_s, which depends on the preference state u_s as determined by the ResNets.
- $e^{-\beta(s-t)}$ is a discounting factor where $\beta > 0$ denotes the discount rate.
- $\mathcal{L}(u_s,\theta_s)$ is the regularization term representing the preference adjustment costs, the parameter $\lambda \geq 0$ governs the strength of the adjustment cost.
- The control variable θ_s represents the ResNets' parameters at time s.

The model employs a mean-field approach, wherein the distribution of ex-ante optimized preference parameters, θ_s, is optimized rather than individual parameters. This formulation, motivated by the high-dimensionality of the parameter space, is analogous to a mean-field optimal control problem, where the objective function seeks to minimize structural risk as in deep learning contexts (e.g., Weinan et al. (2019)). We optimize the distribution of parameters over time, seeking an emergent preference representation that best fits the observed consumption behavior and maximizes discounted utility.

3. NEURAL DIFFERENTIAL UTILITY

This section introduces the Neural Differential Utility (NDU) framework. We derive NDU as a continuous-time limit of a discrete-time recursive utility model and present the formal definition.

3.1. *NDU as a Limit of Discrete-time Recursive Utility*

Consider a discrete-time economy with time periods $t = 0, \Delta t, 2\Delta t, ..., T$. The agent's utility at time t is:

$$V_t = \mathbb{E}_t\left[V_{t+\Delta t} + \Delta t \cdot f_{\theta_t}(c_t, u_t, V_{t+\Delta t})\right], \tag{14}$$

where c_t is consumption, u_t is the preference state, and f_{θ_t} is an aggregator function. The preference state evolves as:

$$u_{t+\Delta t} = u_t + \Delta t \cdot h_{\theta_t}(u_t, c_t), \tag{15}$$

where h_{θ_t} is a ResNets residual block. The cost of preference adjustment is:

$$\frac{\lambda}{2} \sum_{s=t}^{T-\Delta t} \Delta t \|h_{\theta_s}(u_s, c_s)\|^2. \tag{16}$$

The agent maximizes the recursive utility function, net of the preference adjustment cost:

$$\sup_{\theta_s} \mathbb{E}_t \left[\sum_{s=t}^{T-\Delta t} \left(\Delta t \cdot f_{\theta_s}(c_s, u_s, V_{s+\Delta t}) - \frac{\lambda \Delta t}{2} \|h_{\theta_s}(u_s, c_s)\|^2 \right) + V_T \right]. \tag{17}$$

We derive the expression for V_t by iteratively expanding equation (14), using the law of iterated expectations:

$$V_t = \mathbb{E}_t \left[\sum_{s=t}^{T-\Delta t} \Delta t \cdot f_{\theta_s}(c_s, u_s, V_{s+\Delta t}) + V_T \right].$$

We can rewrite equation (17) using dynamic programming, and applying this recursively with $V_T(u_T) = 0$ as:

$$\begin{aligned} V_t(u_t) &= \sup_{\theta} \mathbb{E}_t \left[V_{t+\Delta t}(u_{t+\Delta t}) + \Delta t \cdot f_{\theta_t}(c_t, u_t, V_{t+\Delta t}(u_{t+\Delta t})) - \frac{\lambda \Delta t}{2} \|h_{\theta_t}(u_t, c_t)\|^2 \right] \\ &= \sup_{\theta} \mathbb{E}_t \left[\sum_{s=t}^{T-\Delta t} \left(\frac{V_{s+\Delta t}(u_{s+\Delta t}) - V_s(u_s)}{\Delta t} + f_{\theta_s}(c_s, u_s, V_{s+\Delta t}(u_{s+\Delta t})) \right. \right. \\ &\quad \left. \left. - \frac{\lambda}{2} \|h_{\theta_s}(u_s, c_s)\|^2 \right) \Delta t \right]. \end{aligned}$$

Note that the cost function is rescaled by Δt to ensure a well-defined continuous-time limit. To derive the continuous-time limit, we make the following assumptions:

Assumption 1—Convergence of Processes: *As $\Delta t \to 0$, c_t, u_t, θ_t, and V_t converge to continuous-time processes $c(t)$, $u(t)$, $\theta(t)$, and $V(t)$ in distribution.*

Assumption 2—Convergence of Functions: *As $\Delta t \to 0$, f_{θ_t} and h_{θ_t} converge uniformly to $f_{\theta(t)}$ and $h_{\theta(t)}$.*

Assumption 3—Convergence of Expectation: *As $\Delta t \to 0$, $\mathbb{E}_t[\cdot]$ converges to $\mathbb{E}[\cdot|\mathcal{F}_t]$.*

Assumption 4—Regularity and BSDE Representation: *V_t converges to a sufficiently regular $V(t)$, specifically $V \in C^{1,2}([0,T] \times \mathbb{R}^q \times \mathcal{C} \times \Theta)$. f_θ and h_θ are Lipschitz continuous. The derivatives of V satisfy polynomial growth conditions. There exists a viscosity solution to the HJB equation associated with the continuous-time optimization problem. The solution $V(t)$ to the HJB equation can be represented as the solution to a BSDE. As $\Delta t \to 0$, the difference quotient $(V_{t+\Delta t} - V_t)/\Delta t$ converges in probability to the solution of the BSDE, uniformly in $t \in [0,T]$.*[23]

Under these assumptions, we take the limit as $\Delta t \to 0$ in (15) and (3.1) to obtain the continuous-time model. From (15), we get Equation (Forward Propagation). From (3.1), the summation $\sum_{s=t}^{T-\Delta t} \left(\frac{V_{s+\Delta t}(u_{s+\Delta t}) - V_s(u_s)}{\Delta t} \right) \Delta t$ now converges to $\int_t^T dV_s = V_T - V_t$. Since we have the extra V_T term in (3.1) and we set $V_T = 0$, we obtain $-V_t$. Taking the limit as $\Delta t \to 0$, applying assumption 4, we get the Equation (Backward Propagation). In particular, we have replaced the cost term $\frac{\lambda}{2}\|h_{\theta_s}(u_s, c_s)\|^2$ with a general cost function $\lambda \mathcal{L}(u_s, \theta_s)$. See Definition 4 for the continuous-time model.

Definition 4—Neural Differential Utility: Consider an agent with a finite time horizon $T > 0$ operating within a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$. The consumption process $c = \{c_t\}_{t\in[0,T]}$, where $c_t \in \mathcal{C}_t \subset \mathbb{R}^n$ and $\mathcal{C}_t$ is a closed, convex, and time-dependent set. We assume that c is progressively measurable with

[23] This assumption is crucial for the validity of the continuous-time limit. The Lipschitz and linear growth conditions on f_θ and h_θ are standard in the theory of BSDEs and ensure that the BSDE has a unique solution (see, e.g., Pardoux and Peng (1990)). The convergence of the difference quotient to the BSDE solution can be established under these conditions using results on the convergence of numerical schemes for BSDEs. For instance, Zhang (2004) provides conditions for the convergence of the Euler scheme for BSDEs, which is closely related to our discrete-time approximation. Further, the regularity of $V(t)$ ensures that the Itô calculus can be applied to derive the HJB equation.

respect to $\{\mathcal{F}_t\}$ and that $\mathbb{E}[\int_0^T \|c_t\|^2 dt] < \infty$. The agent also chooses an $\mathcal{F}_t$-adapted control process $\theta = \{\theta_t\}_{t\in[0,T]}$, where $\theta_t \in \Theta \subset \mathbb{R}^{d_\theta}$ and Θ is a convex, compact set. Let $\mathcal{A}$ denote the space of admissible control processes, i.e., progressively measurable processes with values in Θ, such that $\mathbb{E}\left[\int_0^T \|\theta_t\|^2 dt\right] < \infty$. The agent's continuation utility at time s, denoted by $V_s \in \mathbb{R}$, and her dynamic preference state, $u_t \in \mathbb{R}^q$, are governed by the following FBSDE system:

$$du_t = h_{\theta_t}(u_t, c_t)dt + \sigma^h_{\theta_t}(u_t, c_t)d\mathcal{Z}^h_t, \quad u_0 = \phi(c_0), \qquad \text{(Forward Propagation)}$$

$$dV_s = -[f_{\theta_s}(c_s, u_s, V_s) - \lambda\mathcal{L}(u_s, \theta_s)]ds + \sigma_s d\mathcal{Z}_s, \quad V_T = 0, \qquad \text{(Backward Propagation)}$$

where $\theta_t = (A_t, B_t, \Gamma_t, b_t)$ with $A_t \in \mathbb{R}^{q\times m}$, $B_t \in \mathbb{R}^{m\times q}$, $b_t \in \mathbb{R}^m$, and $\Gamma \in \mathbb{R}^{m\times n}$. The function $h_{\theta_t}(u_t, c_t)$ is defined in (12). $\mathcal{Z}_t$ and $\mathcal{Z}^h_t$ are square-integrable $\{\mathcal{F}_t\}$-martingales, σ_t and $\sigma^h_{\theta_t}$ are adapted processes. The other components are:

- $f_{\theta_t}(c_t, u_t, V_t)$: The aggregator function, parameterized by ResNets. We assume f is Lipschitz continuous in (c, u, V) for each fixed θ and continuous in θ for fixed (c, u, V).
- $\phi : \mathcal{C}_0 \to \mathbb{R}^q$: A Lipschitz continuous function mapping the initial consumption c_0 to the initial preference state u_0.
- $\lambda \geq 0$: A scaling parameter representing the weight of cognitive and affective preference adjustment costs.
- $\mathcal{L}(u_t, \theta_t)$: The cognitive and affective preference adjustment costs function. We primarily focus on a quadratic cost function: $\mathcal{L}(u_t, \theta_t) = \|\theta_t - \theta_0\|^2$, where θ_0 represents a baseline or default control strategy.[24]

Optimization Problem. The agent's optimization problem at time t is given by:

$$V_t(c;\theta) = \sup_{\theta_s\in\mathcal{A}} \mathbb{E}\left[\int_t^T \left(f_{\theta_s}(c_s, u_s, V_s) - \lambda\mathcal{L}(u_s, \theta_s)\right) ds \middle| \mathcal{F}_t\right], \tag{18}$$

[24]This baseline reflects the agent's preference evolution in the absence of active intervention. The quadratic cost function penalizes deviations from this baseline, with larger deviations incurring higher costs. This form captures the concept of increasing marginal cost while maintaining analytical tractability.

subject to (Forward Propagation) and (Backward Propagation), a budget constraint.

Remark 3—NDU, Linear Case, and Habit Formation: NDU's forward propagation, $du_t = A_t(B_t u_t + \Gamma_t c_t)dt + \sigma^h_{\theta_t} d\mathcal{Z}^h_t$, can exhibit dynamics similar to habit formation in a simplified linear setting. Let $u_t = \theta_t c_t$, where $\theta_t = (A_t, B_t, \Gamma_t)$, $b_t = 0$, with exogenous dynamics $d\theta_t = \alpha_t dt$, and consumption follows $dc_t = \mu c_t dt + \sigma c_t d\mathcal{Z}_t$. Assume the dimension of θ_t is $q \times n$, making the dimensions of c_t and u_t consistent. Assume the dimensions of A_t, B_t, and Γ_t are $q \times m$, $m \times q$, and $m \times n$, respectively, to ensure that matrix multiplication is well-defined. Assuming $d\mathcal{Z}^h_t = d\mathcal{Z}_t$ and $d[\theta, c]_t = 0$, Itô's lemma applied to $u_t = \theta_t c_t$, yields:

$$du_t = d(\theta_t c_t) = c_t d\theta_t + \theta_t dc_t + d[\theta, c]_t = (\alpha_t c_t + \mu \theta_t c_t)dt + \sigma \theta_t c_t d\mathcal{Z}_t,$$

where $\sigma^h_{\theta_t} = \sigma c_t \theta_t$, linking preference volatility to consumption volatility and ResNets parameters. A system of equations: $A_t(B_t\theta_t + \Gamma_t) = \alpha_t + \mu\theta_t$ (for $c_t \neq 0$), relating parameter dynamics (α), consumption growth (μ), and ResNets parameters (A_t, B_t, Γ_t). These results demonstrate that a simplified linear NDU can encompass habit formation when controlled parameters are exogenously given (e.g., $\theta_t = 1/h_t$).

3.2. *Existence and Uniqueness of Utility*

To ensure the well-posedness of the model, we need to establish the existence and uniqueness of a solution to the FBSDE system (Forward Propagation)-(Backward Propagation) for a given control process θ, as well as the existence of an optimal control process θ^* that solves the agent's optimization problem (18). We make the following assumptions, which are adapted from Duffie and Epstein (1992) and Yong and Zhou (2012) to our setting:

Assumption 5—Regularity Conditions for Existence and Uniqueness: *The functions h_{θ_t}, f_{θ_t}, $\mathcal{L}$, $\sigma^h_{\theta_t}$, and ϕ satisfy the following conditions:*[25]

[25]The Lipschitz conditions in Assumption 5 have important economic interpretations. They ensure the model's well-behavedness by preventing overly sensitive or erratic responses in the agent's preference state, instantaneous utility, and preference adjustment cost to small changes in consumption, the

(i) $h_{\theta_t}(u,c)$ *is measurable in* (t,θ,u,c) *and satisfies a Lipschitz condition in* (u,c) *uniformly in* θ *and* t*. That is, there exists a constant* $L_h > 0$ *such that for all* $u,u',c,c' \in \mathbb{R}^m$ *and* $\theta \in \Theta$, $t \in [0,T]$*:*

$$\|h_{\theta_t}(u,c) - h_{\theta_t}(u',c')\| \leq L_h(\|u-u'\| + \|c-c'\|).$$

(ii) $f_{\theta_t}(c,u,V)$ *is measurable in* (t,θ,c,u,V) *and satisfies a Lipschitz condition in* (c,u,V) *uniformly in* θ *and* t*. That is, there exists a constant* $L_f > 0$ *such that for all* $c,c' \in \mathcal{C}$*,* $u,u' \in \mathbb{R}^q$*,* $V,V' \in \mathbb{R}$*, and* $\theta \in \Theta$, $t \in [0,T]$*:*

$$\|f_{\theta_t}(c,u,V) - f_{\theta_t}(c',u',V')\| \leq L_f(\|c-c'\| + \|u-u'\| + \|V-V'\|).$$

(iii) The preference adjustment cost function $\mathcal{L}(u,\theta)$ *is measurable in* (u,θ) *and satisfies a Lipschitz condition. There exists a constant* $L_{\mathcal{L}} > 0$ *such that for all* $u,u' \in \mathbb{R}^q$ *and* $\theta,\theta' \in \Theta$*:*

$$|\mathcal{L}(u,\theta) - \mathcal{L}(u',\theta')| \leq L_{\mathcal{L}}(\|u-u'\| + \|\theta-\theta'\|).$$

(iv) $\sigma^h_{\theta_t}$ *is Lipschitz continuous in* u *and* c*, uniformly in* t*. That is, there exists a constant* $L^h_\sigma > 0$ *such that for all* $t \in [0,T]$*,* $u,u' \in \mathbb{R}^q$*, and* $c,c' \in \mathbb{R}^n$*:*

$$\|\sigma^h_{\theta_t}(u,c) - \sigma^h_{\theta_t}(u',c')\| \leq L_{\sigma^h}(\|u-u'\| + \|c-c'\|).$$

(v) $\phi : \mathcal{C}_0 \to \mathbb{R}^q$ *is Lipschitz continuous, i.e., there exists a constant* $L_\phi > 0$ *such that for all* $c_0, c_0' \in \mathcal{C}_0$*:*

$$\|\phi(c_0) - \phi(c_0')\| \leq L_\phi \|c_0 - c_0'\|.$$

Assumption 6—Growth Conditions: *The functions* h_{θ_t}*,* f_{θ_t}*,* $\mathcal{L}$*, and* $\sigma^h_{\theta_t}$ *satisfy the following linear and quadratic growth conditions. There exist constants* $K_h, K_f, K_{\mathcal{L}}, K_{\sigma^h} > 0$

---

preference state, continuation utility, or the model parameters. Specifically, the Lipschitz conditions on h_{θ_t}, f_{θ_t}, $\mathcal{L}$, $\sigma^h_{\theta_t}$, and ϕ prevent, respectively, abrupt changes in the drift of the preference state, instantaneous utility, initial preference state, and preference adjustment cost in response to small changes in their respective arguments. These conditions are crucial for the existence and uniqueness of solutions to the FBSDE system.

such that for all $u \in \mathbb{R}^q$, $c \in \mathcal{C}$, $V \in \mathbb{R}$, *and* $\theta \in \Theta$, $t \in [0,T]$:[26]

$$\|h_{\theta_t}(u,c)\| \leq K_h(1 + \|u\| + \|c\|),$$
$$\|f_{\theta_t}(c,u,V)\| \leq K_f(1 + \|c\| + \|u\| + \|V\|),$$
$$|\mathcal{L}(u,\theta)| \leq K_{\mathcal{L}}(1 + \|u\|^2 + \|\theta\|^2),$$
$$\|\sigma^h_{\theta_t}(u,c)\| \leq K_{\sigma^h}(1 + \|u\| + \|c\|).$$

ASSUMPTION 7—Conditions for Optimal Control: *The set* Θ *of admissible controls is a convex and compact subset of a finite-dimensional Euclidean space. The function* $H(t,c,u,V,\theta) = f_\theta(c,u,V) - \lambda\mathcal{L}(u,\theta)$ *is concave in* θ *for each* (t,c,u,V).[27]

THEOREM 3—Existence and Uniqueness: *Let* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ *be a filtered probability space satisfying the usual conditions. Let* $\theta = \{\theta_t\}_{t\in[0,T]}$ *be an adapted control process such that* $\mathbb{E}\left[\int_0^T \|\theta_t\|^2 dt\right] < \infty$, *and let* $c = \{c_t\}_{t\in[0,T]}$ *be a progressively measurable consumption process such that* $\mathbb{E}\left[\int_0^T \|c_t\|^2 dt\right] < \infty$. *Under Assumptions 5, 6, and 7, for any given adapted control process* θ *and progressively measurable consumption process* c *satisfying the stated integrability conditions, there exists a unique pair of adapted processes* $(u,V) = \{(u_t,V_t)\}_{t\in[0,T]}$ *that solves the FBSDE system* (Forward Propagation)-(Backward Propagation), *such that* u_t *is RCLL (right-continuous with left limits) and* $\mathbb{E}\left[\int_0^T (\|u_t\|^2 + |V_t|^2) dt\right] < \infty$. *Moreover, under Assumption 7, there exists an optimal adapted control process* $\theta^* = \{\theta^*_t\}_{t\in[0,T]}$ *that maximizes the agent's objective* (18).

4. PROPERTIES OF THE VALUE FUNCTION

We now analyze the properties of the value function V_t defined in (18).

[26] The growth conditions in Assumption 6 have economic interpretations related to preventing unrealistic behavior in the model. Specifically, they ensure: (1) the drift of the preference state (h_{θ_t}) doesn't explode; (2) instantaneous utility (f_{θ_t}) remains finite; and (3) the preference adjustment cost ($\mathcal{L}$) doesn't become infinite. These conditions are crucial for the well-behavedness of the model and the existence and uniqueness of solutions to the FBSDE system.

[27] The assumptions on Θ (compact and convex) and H (concave in θ) ensure the existence of an optimal control and allow for the application of tools like Pontryagin's maximum principle.

PROPOSITION 4—Continuity: *Under Assumptions 5 and 6, the value function $V_t(c;\theta)$ is continuous in c and θ in the respective L^2 spaces. Specifically, for any $t \in [0,T]$, if $c^n \to c$ in $L^2(\Omega, \mathcal{F}_t; \mathbb{R}^n)$ and $\theta^n \to \theta$ in $L^2([t,T] \times \Omega; \Theta)$, then $V_t(c^n;\theta^n) \to V_t(c;\theta)$ in $L^2(\Omega, \mathcal{F}_t; \mathbb{R})$.*

PROPOSITION 5—Monotonicity: *Assume that $f_{\theta_t}(c_t, u_t, V_t)$ is non-decreasing in c_t for any given u_t, V_t, and θ_t. If $c_t^1 \geq c_t^2$ for all $t \in [0,T]$ (component-wise for vector-valued consumption), and if $\phi(c_0^1) \geq \phi(c_0^2)$, then the value function is non-decreasing in the consumption process, i.e., $V_t(c^1;\theta) \geq V_t(c^2;\theta)$ for fixed θ. This will hold if h is increasing in c and f is increasing in u.*

PROPOSITION 6—Concavity: *Suppose that the aggregator function $f_{\theta_t}(c_t, u_t, V_t)$ is concave in (c_t, V_t) for any given u_t and θ_t, and suppose the activation function η in the definition of h_{θ_t} is linear. Then, if Θ is a convex set, the value function $V_t(c;\theta)$ is concave in c for fixed θ.*

PROPOSITION 7—Dynamic Consistency: *Let (c^*, θ^*) be the joint optimal process starting from time t, solving*

$$V_t = \sup_{c,\theta} \mathbb{E}_t \left[\int_t^T \left(f_{\theta_s}(c_s, u_s, V_s) - \lambda \mathcal{L}(u_s, \theta_s) \right) ds \middle| \mathcal{F}_t \right], \tag{19}$$

subject to (Forward Propagation) *and* (Backward Propagation)*, as well as any relevant constraints on consumption. Then, for any $s \in (t,T]$, the continuation of (c^*, θ^*) from time s onwards is also optimal for the problem starting at time s:*

$$V_s = \sup_{c,\theta} \mathbb{E}_s \left[\int_s^T \left(f_{\theta_\tau}(c_\tau, u_\tau, V_\tau) - \lambda \mathcal{L}(u_\tau, \theta_\tau) \right) d\tau \middle| \mathcal{F}_s \right]. \tag{20}$$

4.1. *Preference Axioms*

We now examine the implications of the NDU framework for the standard axioms of preference relations. We define the preference relation $\succsim_t$ over consumption processes in $\mathcal{D}$ as follows:

DEFINITION 5—Preference Relation: For any two consumption processes $c, c' \in \mathcal{D}$, we say that $c \succsim_t c'$ if and only if $V_t^*(c) \geq V_t^*(c')$, where $V_t^*(c)$ denotes the optimal value function at time t given consumption process c. The strict preference relation $c \succ_t c'$ is defined by $V_t^*(c) > V_t^*(c')$, and the indifference relation $c \sim_t c'$ is defined by $V_t^*(c) = V_t^*(c')$.

Since the preference relation is defined through a real-valued function, the properties of completeness and transitivity follow directly.

AXIOM 1—Completeness and Transitivity: *The preference relation $\succsim_t$ is complete and transitive.*

ASSUMPTION 8: *The aggregator function f_{θ_t} is monotone in c. Specifically, for any c, c' such that $c \geq c'$ (component-wise), and for any fixed u and V, we have $f_{\theta_t}(c, u, V) \geq f_{\theta_t}(c', u, V)$. The residual block function h_{θ_t} is non-decreasing in c.*

AXIOM 2—Monotonicity: *If the aggregator function f_{θ_t} is monotone with respect to c for all θ_t (Assumption 8), and if h_{θ_t} is non-decreasing in c, then the preference relation $\succsim_t$ is monotone. In other words, if $c_s \geq c'_s$ for all $s \in [t, T]$, then $c \succsim_t c'$. If, in addition, there exists a time interval $I \subseteq [t, T]$ with positive measure such that $c_s > c'_s$ for all $s \in I$, and f_{θ_t} is strictly increasing in c over that interval, then $c \succ_t c'$.*

AXIOM 3—Continuity: *Let $\mathcal{D}$ be the space of progressively measurable consumption processes $c = \{c_t\}_{t \in [0,T]}$ such that $\mathbb{E}[\int_0^T \|c_t\|^2 dt] < \infty$, equipped with the L^2 norm $\|c\|_{\mathcal{D}} = (\mathbb{E}[\int_0^T \|c_t\|^2 dt])^{1/2}$. If the aggregator function f_{θ_t}, the terminal utility, and ϕ are continuous, then the preference relation $\succsim_t$ is continuous with respect to the L^2 norm on $\mathcal{D}$.*

4.2. *Implication: A Neuroeconomic Perspective on AI*

NDU bridges AI and neuroeconomics by incorporating resource constraints inherent in human cognition (e.g., Shenhav et al. (2013, 2016)). This contrasts with traditional AI models that often neglect such constraints. By introducing preference adjustment costs and resource-dependent utility, NDU suggests a path towards resource-aware AI that could exhibit more robust and adaptable decision-making,

analogous to human behavior under cognitive limitations. Integrating neuroeconomic principles into AI development, as exemplified by NDU, may be crucial for creating truly intelligent systems that can effectively collaborate with and serve human needs. This approach could lead to more human-compatible AI, capable of understanding and responding to human limitations.

5. CONCLUDING REMARKS

Since (e.g., Allais (1953), Ellsberg (1961), Kahneman and Tversky (1979)), the standard neoclassical framework, particularly its reliance on expected utility theory and rational expectations, has faced challenges. The ongoing debate between behavioral and neoclassical approaches underscores the need for a more rigorous and generalizable theoretical framework for behavioral economics. The recent advancements in deep learning for directly modeling neural functioning and reinforcement learning for modeling dynamic decision-making provide new insights for developing such a framework. These methods offer the potential to develop more unified models of behavior results and to analyze the impact of cognitive and affective biases on market outcomes.

This paper attempts to formalize the two-dimensional framework (see Table I) within NDU, building on existing literature (e.g., Camerer et al. (2005), Kahneman (2011)), and proves the universal approximation theorem within this framework (see Qi (2025d)). Although we cannot conclude NDU encompass all behavioral results, particularly those beyond the L^p spaces, it offers a feasible strategy to bridge the gap between disparate behavioral models and the neoclassical framework by incorporating resource constraint friction (e.g., Gershman et al. (2015), Lieder and Griffiths (2020), Bhui et al. (2021)). We also present four representative cases demonstrating how a resource-constrained rational agent can endogenously behave in risk averse, ambiguity averse, loss averse, and present bias (see Qi (2025c,a,b)).

The NDU framework also offers a valuable perspective for designing artificial intelligence (AI) systems by emphasizing the evolution of preferences under resource constraints. By modeling dynamic adaptation to environmental feedback and computational costs, NDU helps understand how AIs develop heuristics, simplify models, and shift priorities due to resource limitations. This contributes to

interpretable AI, enabling the design of more robust, ethical, and trustworthy AI systems that adapt to changing environments.

APPENDIX A: MAIN PROOFS

PROOF OF PROPOSITION 1: The HJB equation for the agent's problem is given by (3). The first-order condition yields $\theta^*(c) = 1/\lambda$. Substituting into (3) yields $0 = \frac{c}{2\lambda} - \beta V(c) + \mu c V'(c) + \frac{1}{2}\sigma^2 c^2 V''(c)$. Postulating a solution of the form $V(c) = Ac + Bc^\gamma$, and substituting it into the equation above, the coefficients of c and c^γ must vanish. This implies $A = \frac{1}{2\lambda(\beta-\mu)}$ (given $\beta > \mu$) and that γ solves the quadratic equation $\frac{1}{2}\sigma^2\gamma^2 + (\mu - \frac{1}{2}\sigma^2)\gamma - \beta = 0$ We choose the negative root, $\gamma = \frac{-(\mu-\frac{1}{2}\sigma^2)-\sqrt{(\mu-\frac{1}{2}\sigma^2)^2+2\sigma^2\beta}}{\sigma^2} < 0$, to ensure the transversality condition (TVC) is satisfied. The boundary condition $V(1) = 0$ implies $B = -A$. Thus, $V(c) = \frac{c-c^\gamma}{2\lambda(\beta-\mu)}$. The coefficient of relative risk aversion is $\mathcal{R}(c) = -\frac{cV''(c)}{V'(c)} = \frac{\gamma(\gamma-1)c^{\gamma-1}}{1-\gamma c^{\gamma-1}}$. Since $\gamma < 0$, $\lim_{c\to 0}\mathcal{R}(c) = 1-\gamma > 1$ and $\lim_{c\to\infty}\mathcal{R}(c) = 0$. *Q.E.D.*

PROOF OF THEOREM 3: The proof consists of two parts: (1) existence and uniqueness of the solution to the FBSDE (Forward Propagation)-(Backward Propagation) system for a given control process θ, and (2) existence of an optimal control process θ^* that solves the agent's optimization problem (18).

Part 1: Existence and Uniqueness of the FBSDE Solution.

Step 1: Existence and Uniqueness of u_t for a given control θ_t. See Proof of Proposition 2 in Supplement.

Step 2: Existence and Uniqueness of V_t for a given control θ_t. Given the existence of an RCLL process u_t established in Step 1, we can define the process $Z_t = (c_t, u_t, \theta_t)$ and rewrite the backward propagation equation (Backward Propagation) as:

$$V_t = \mathbb{E}_t\left[\int_t^T \left(f_{\theta_s}(Z_s, V_s) - \lambda\mathcal{L}(Z_s)\right) ds\right]. \tag{21}$$

This is a recursive equation for V_t of the form considered in Duffie and Epstein (1992). We can apply Proposition A1 from Appendix A of Duffie and Epstein (1992) to establish the existence and uniqueness of a solution V_t. To apply Proposition A1, we need to verify that the following conditions hold:

(i) The function $(z, v) \mapsto f_{\theta_t}(z, v) - \lambda \mathcal{L}(z)$ is measurable, uniformly Lipschitz in v, and uniformly bounded when v is bounded.

(ii) The process $Z_t = (c_t, u_t, \theta_t)$ is progressively measurable and has finite second moments, i.e., $\mathbb{E}[\int_0^T \|Z_t\|^2 dt] < \infty$.

Verification of Condition (i):

- **Measurability**: The function $(z, v) \mapsto f_{\theta_t}(z, v) - \lambda \mathcal{L}(z)$ is measurable because f_{θ_t} and $\mathcal{L}$ are measurable by Assumption 5.
- **Uniform Lipschitz Continuity in** v: By Assumption 5 (ii), f_{θ_t} is Lipschitz continuous in (c, u, V) uniformly in θ and t. This implies that f_{θ_t} is Lipschitz continuous in V uniformly in $Z = (c, u, \theta)$ and t. Specifically, there exists a constant $L_f > 0$ such that for all $z \in \mathcal{C} \times \mathbb{R}^q \times \Theta$, $v, v' \in \mathbb{R}$, and $t \in [0, T]$:

$$\|f_{\theta_t}(z, v) - f_{\theta_t}(z, v')\| \leq L_f |v - v'|.$$

  Since $\mathcal{L}$ does not depend on v, the function $(z, v) \mapsto f_{\theta_t}(z, v) - \lambda \mathcal{L}(z)$ is also uniformly Lipschitz in v.
- **Uniform Boundedness when** v **is bounded**: Suppose $|v| \leq K$ for some constant $K > 0$. By Assumption 6, f_{θ_t} satisfies a linear growth condition, and $\mathcal{L}$ satisfies a quadratic growth condition. Since c_t and u_t have finite second moments (established in Step 1), and θ_t is bounded, the process $Z_t = (c_t, u_t, \theta_t)$ is bounded in expectation. Therefore, there exists a constant $M > 0$ such that for all $t \in [0, T]$ and $|v| \leq K$:

$$\mathbb{E}\left[|f_{\theta_t}(Z_t, v) - \lambda \mathcal{L}(Z_t)|\right] \leq M.$$

  This implies that $f_{\theta_t}(Z_t, v) - \lambda \mathcal{L}(Z_t)$ is uniformly bounded in expectation when v is bounded.

Verification of Condition (ii):

- **Progressive Measurability**: The process $Z_t = (c_t, u_t, \theta_t)$ is progressively measurable because c_t and θ_t are progressively measurable by assumption, and u_t

is the solution to an SDE driven by a progressively measurable process, which makes u_t also progressively measurable.

- **Finite Second Moments**: We have $\mathbb{E}[\int_0^T \|c_t\|^2 dt] < \infty$ and $\mathbb{E}[\int_0^T \|\theta_t\|^2 dt] < \infty$ by assumption. From Step 1, we have $\mathbb{E}[\sup_{0\le t\le T} \|u_t\|^2] < \infty$, which implies $\mathbb{E}[\int_0^T \|u_t\|^2 dt] < \infty$. Therefore, $\mathbb{E}[\int_0^T \|Z_t\|^2 dt] = \mathbb{E}[\int_0^T (\|c_t\|^2 + \|u_t\|^2 + \|\theta_t\|^2) dt] < \infty$.

Since both conditions (i) and (ii) are satisfied, we can apply Proposition A1 from Duffie and Epstein (1992). This proposition guarantees the existence of a unique progressively measurable process V_t such that (21) holds, and $\mathbb{E}[\int_0^T |V_t|^2 dt] < \infty$. Moreover, V_t has right-continuous paths with left limits. Furthermore, by the same proposition, there exists a progressively measurable process σ_t such that $\mathbb{E}[\int_0^T \|\sigma_t\|^2 dt] < \infty$ and

$$V_t = V_0 - \int_0^t (f_{\theta_s}(Z_s, V_s) - \lambda \mathcal{L}(Z_s)) ds + \int_0^t \sigma_s d\mathcal{Z}_s, \quad V_T = 0.$$

This is equivalent to the BSDE (Backward Propagation). Thus, we have established the existence and uniqueness of a solution (V_t, σ_t) to the BSDE (also see Pardoux and Peng (1990)) for a given control θ_t, with the σ_t as an auxiliary process that is implicitly determined by the BSDE.

Step 3: Coupling the Forward and Backward SDEs We define the following iterative scheme for $n \ge 0$:

$$\begin{aligned} du_t^{n+1} &= h_{\theta_t}(u_t^n, c_t) dt + \sigma_{\theta_t}^h(u_t^n, c_t) d\mathcal{Z}_t^h, \quad u_0^{n+1} = \phi(c_0), \\ dV_t^{n+1} &= -[f_{\theta_t}(c_t, u_t^{n+1}, V_t^{n+1}) - \lambda \mathcal{L}(u_t^{n+1}, \theta_t)] dt + \sigma_t^{n+1} d\mathcal{Z}_t, \quad V_T^{n+1} = 0. \end{aligned}$$

We start with $u_t^0 = \phi(c_0)$. By Steps 1 and 2, for each n, there exist unique solutions u_t^{n+1} and $(V_t^{n+1}, \sigma_t^{n+1})$. We need to show that $(u_t^n, V_t^n, \sigma_t^n)$ converges to a limit (u_t, V_t, σ_t) in an appropriate sense, and that this limit is the unique solution to the coupled FBSDE system (Forward Propagation)-(Backward Propagation). Let

$\delta u_t^n = u_t^{n+1} - u_t^n$, $\delta V_t^n = V_t^{n+1} - V_t^n$, and $\delta \sigma_t^n = \sigma_t^{n+1} - \sigma_t^n$. Then, we have:

$$
\begin{aligned}
d(\delta u_t^n) &= [h_{\theta_t}(u_t^n, c_t) - h_{\theta_t}(u_t^{n-1}, c_t)]dt + [\sigma_{\theta_t}^h(u_t^n, c_t) - \sigma_{\theta_t}^h(u_t^{n-1}, c_t)]d\mathcal{Z}_t^h, \\
d(\delta V_t^n) &= -[f_{\theta_t}(c_t, u_t^{n+1}, V_t^{n+1}) - f_{\theta_t}(c_t, u_t^n, V_t^n) - \lambda(\mathcal{L}(u_t^{n+1}, \theta_t) - \mathcal{L}(u_t^n, \theta_t))]dt \\
&\quad + \delta \sigma_t^n d\mathcal{Z}_t.
\end{aligned}
$$

Using the Lipschitz continuity of h_{θ_t}, f_{θ_t}, $\sigma_{\theta_t}^h$, and $\mathcal{L}$, applying Itô's formula to $\|\delta u_t^n\|^2$ and $|\delta V_t^n|^2$, and taking expectation, we obtain the following estimates.

Detailed Derivation of the Estimates: Applying Itô's formula to $\|\delta u_t^n\|^2$, we get:

$$
\begin{aligned}
d\|\delta u_t^n\|^2 &= 2\langle \delta u_t^n, d(\delta u_t^n)\rangle + \|\sigma_{\theta_t}^h(u_t^n, c_t) - \sigma_{\theta_t}^h(u_t^{n-1}, c_t)\|^2 dt \\
&= 2\langle \delta u_t^n, h_{\theta_t}(u_t^n, c_t) - h_{\theta_t}(u_t^{n-1}, c_t)\rangle dt + 2\langle \delta u_t^n, [\sigma_{\theta_t}^h(u_t^n, c_t) - \sigma_{\theta_t}^h(u_t^{n-1}, c_t)]d\mathcal{Z}_t^h\rangle \\
&\quad + \|\sigma_{\theta_t}^h(u_t^n, c_t) - \sigma_{\theta_t}^h(u_t^{n-1}, c_t)\|^2 dt.
\end{aligned}
$$

Taking expectation and using the Lipschitz continuity of h and σ^h, we have:

$$
\begin{aligned}
\frac{d}{dt}\mathbb{E}[\|\delta u_t^n\|^2] &= 2\mathbb{E}\left[\langle \delta u_t^n, h_{\theta_t}(u_t^n, c_t) - h_{\theta_t}(u_t^{n-1}, c_t)\rangle\right] + \mathbb{E}\left[\|\sigma_{\theta_t}^h(u_t^n, c_t) - \sigma_{\theta_t}^h(u_t^{n-1}, c_t)\|^2\right] \\
&\leq 2\mathbb{E}\left[\|\delta u_t^n\| L_h \|\delta u_t^{n-1}\|\right] + L_{\sigma^h}^2 \mathbb{E}\left[\|\delta u_t^{n-1}\|^2\right] \\
&\leq L_h \mathbb{E}\left[\|\delta u_t^n\|^2 + \|\delta u_t^{n-1}\|^2\right] + L_{\sigma^h}^2 \mathbb{E}\left[\|\delta u_t^{n-1}\|^2\right] \quad (\text{using } 2ab \leq a^2 + b^2) \\
&= L_h \mathbb{E}[\|\delta u_t^n\|^2] + (L_h + L_{\sigma^h}^2)\mathbb{E}[\|\delta u_t^{n-1}\|^2].
\end{aligned}
$$

Let $y_n(t) = \mathbb{E}[\|\delta u_t^n\|^2]$. Then the above inequality becomes

$$
\frac{d}{dt} y_n(t) \leq L_h y_n(t) + (L_h + L_{\sigma^h}^2) y_{n-1}(t).
$$

Let $C = L_h + L_{\sigma^h}^2$. Multiplying both sides by $e^{-L_h t}$, we have

$$
\frac{d}{dt}(y_n(t) e^{-L_h t}) \leq C y_{n-1}(t) e^{-L_h t}.
$$

Integrating both sides from 0 to t, we get

$$y_n(t)e^{-L_h t} \le \int_0^t C y_{n-1}(s)e^{-L_h s}ds, \quad \text{thus,} \quad y_n(t) \le \int_0^t C y_{n-1}(s)e^{L_h(t-s)}ds.$$

By iterating this inequality, we have

$$\begin{aligned} y_n(t) &\le \int_0^t Ce^{L_h(t-s_1)} \int_0^{s_1} Ce^{L_h(s_1-s_2)} y_{n-2}(s_2)ds_2 ds_1 \\ &= \int_0^t \int_0^{s_1} C^2 e^{L_h(t-s_2)} y_{n-2}(s_2)ds_2 ds_1 \\ &\le \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} C^n e^{L_h(t-s_n)} y_0(s_n)ds_n \cdots ds_2 ds_1 \\ &\le \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} C^n e^{L_h(t-s_n)} M ds_n \cdots ds_2 ds_1, \end{aligned} \tag{22}$$

where $M = \sup_{0\le t\le T} y_0(t)$. Let $t \ge s_1 \ge \cdots \ge s_n \ge 0$. Then $t - s_n \le t$, and thus $e^{L_h(t-s_n)} \le e^{L_h t}$. Also, we can bound the integral as follows:

$$\int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} ds_n \cdots ds_2 ds_1 = \frac{t^n}{n!}.$$

Therefore, we have

$$y_n(t) \le \frac{(Ct)^n}{n!} e^{L_h t} M \le \frac{(CT)^n}{n!} e^{L_h T} M, \tag{23}$$

Similarly, applying Itô's formula to $|\delta V_t^n|^2$ and taking expectation, we get:

$$\begin{aligned} d|\delta V_t^n|^2 &= 2\delta V_t^n d(\delta V_t^n) + \|\delta\sigma_t^n\|^2 dt \\ &= -2\delta V_t^n [f_{\theta_t}(c_t, u_t^{n+1}, V_t^{n+1}) - f_{\theta_t}(c_t, u_t^n, V_t^n) - \lambda(\mathcal{L}(u_t^{n+1}, \theta_t) - \mathcal{L}(u_t^n, \theta_t))]dt \\ &\quad + 2\delta V_t^n \delta\sigma_t^n d\mathcal{Z}_t + \|\delta\sigma_t^n\|^2 dt. \end{aligned}$$

Taking the expectation and noting that $\mathbb{E}[\int_t^T 2\delta V_s^n \delta\sigma_s^n d\mathcal{Z}_s] = 0$ because it's an integral of a predictable process with respect to a martingale, we have:

$$\begin{aligned}\mathbb{E}[|\delta V_t^n|^2] = \mathbb{E}[|\delta V_T^n|^2] - \mathbb{E}\left[\int_t^T \|\delta\sigma_s^n\|^2 ds\right] \\ + 2\mathbb{E}\left[\int_t^T \delta V_s^n (f_{\theta_s}(c_s, u_s^{n+1}, V_s^{n+1}) - f_{\theta_s}(c_s, u_s^n, V_s^n))ds\right] \\ - 2\lambda\mathbb{E}\left[\int_t^T \delta V_s^n (\mathcal{L}(u_s^{n+1}, \theta_s) - \mathcal{L}(u_s^n, \theta_s))ds\right].\end{aligned}$$

Since $V_T^{n+1} = V_T^n = 0$, we have $\delta V_T^n = 0$. Thus,

$$\begin{aligned}\mathbb{E}[|\delta V_t^n|^2] + \mathbb{E}\left[\int_t^T \|\delta\sigma_s^n\|^2 ds\right] = 2\mathbb{E}\left[\int_t^T \delta V_s^n (f_{\theta_s}(c_s, u_s^{n+1}, V_s^{n+1}) - f_{\theta_s}(c_s, u_s^n, V_s^n))ds\right] \\ - 2\lambda\mathbb{E}\left[\int_t^T \delta V_s^n (\mathcal{L}(u_s^{n+1}, \theta_s) - \mathcal{L}(u_s^n, \theta_s))ds\right].\end{aligned}$$

Now, we use the Lipschitz continuity of f_{θ_t} and $\mathcal{L}$. Recall that f_{θ_t} is Lipschitz in (c, u, V) uniformly in θ and t, and $\mathcal{L}$ is Lipschitz in u uniformly in θ. Let L_f and $L_{\mathcal{L}}$ be the Lipschitz constants. Then, we can bound the right-hand side as follows:

$$\begin{aligned}&2\mathbb{E}\left[\int_t^T \delta V_s^n (f_{\theta_s}(c_s, u_s^{n+1}, V_s^{n+1}) - f_{\theta_s}(c_s, u_s^n, V_s^n))ds\right] \\ &- 2\lambda\mathbb{E}\left[\int_t^T \delta V_s^n (\mathcal{L}(u_s^{n+1}, \theta_s) - \mathcal{L}(u_s^n, \theta_s))ds\right] \\ &\leq 2\mathbb{E}\left[\int_t^T |\delta V_s^n|(L_f\|\delta u_s^n\| + L_f|\delta V_s^n|)ds\right] + 2|\lambda|\mathbb{E}\left[\int_t^T |\delta V_s^n| L_{\mathcal{L}}\|\delta u_s^n\| ds\right] \\ &\leq C_V\mathbb{E}\left[\int_t^T |\delta V_s^n|(\|\delta u_s^n\| + |\delta V_s^n|)ds\right] \quad (\text{where } C_V \text{ depends on } L_f, L_{\mathcal{L}}, \lambda) \\ &\leq C_V\mathbb{E}\left[\int_t^T (|\delta V_s^n|^2 + \|\delta u_s^n\|^2)ds\right] \quad (\text{using } 2ab \leq a^2 + b^2).\end{aligned}$$

Therefore, we have:

$$\mathbb{E}[|\delta V_t^n|^2] + \mathbb{E}\left[\int_t^T \|\delta\sigma_s^n\|^2 ds\right] \leq C_V\mathbb{E}\left[\int_t^T (|\delta V_s^n|^2 + \|\delta u_s^n\|^2)ds\right].$$

Substituting (23) into the inequality for δV_t^n, we get:

$$\begin{aligned}\mathbb{E}[|\delta V_t^n|^2] + \mathbb{E}\left[\int_t^T \|\delta\sigma_s^n\|^2 ds\right] &\leq C_V \int_t^T \mathbb{E}[|\delta V_s^n|^2] ds + C_V \int_t^T \mathbb{E}[\|\delta u_s^n\|^2] ds \\ &\leq C_V \int_t^T \mathbb{E}[|\delta V_s^n|^2] ds + C_V \int_t^T \frac{(Cs)^n}{n!} e^{L_h s} M ds.\end{aligned}$$

Let $z_n(t) = \mathbb{E}[|\delta V_t^n|^2]$. Then, we have

$$z_n(t) \leq C_V \int_t^T z_n(s) ds + C_V \int_t^T \frac{(Cs)^n}{n!} e^{L_h s} M ds.$$

Let $M_V = \sup_{0\leq t\leq T} z_0(t)$. Similar to how we derive equation (22), we have

$$\begin{aligned}z_n(t) &\leq \int_t^T \int_t^{s_1} \cdots \int_t^{s_{n-1}} C_V^n e^{C_V(T-s_n)} \left(C_V \int_t^T \frac{(Cs)^n}{n!} e^{L_h s} M ds\right) ds_n \cdots ds_2 ds_1 \\ &\leq \frac{(C_V(T-t))^n}{n!} e^{C_V(T-t)} C_V \int_t^T \frac{(Cs)^n}{n!} e^{L_h s} M ds.\end{aligned}$$

Thus,

$$\mathbb{E}[|\delta V_t^n|^2] \leq \frac{(C_V T)^n}{n!} e^{C_V T} C_V \int_0^T \frac{(Cs)^n}{n!} e^{L_h s} M ds. \tag{24}$$

Also,

$$\mathbb{E}\left[\int_t^T \|\delta\sigma_s^n\|^2 ds\right] \leq C_V \int_t^T \mathbb{E}[|\delta V_s^n|^2] ds + C_V \int_t^T \frac{(Cs)^n}{n!} e^{L_h s} M ds. \tag{25}$$

Convergence of the Iterative Scheme: From the estimates (23), (24), and (25), we can see that the sequences $\{\mathbb{E}[\|\delta u_t^n\|^2]\}$, $\{\mathbb{E}[|\delta V_t^n|^2]\}$, and $\{\mathbb{E}[\int_t^T \|\delta\sigma_s^n\|^2 ds]\}$ converge to 0 as $n \to \infty$. Specifically, for any $t \in [0, T]$, we have:

$$\lim_{n\to\infty} \mathbb{E}[\|\delta u_t^n\|^2] = 0, \quad \lim_{n\to\infty} \mathbb{E}[|\delta V_t^n|^2] = 0, \quad \lim_{n\to\infty} \mathbb{E}\left[\int_t^T \|\delta\sigma_s^n\|^2 ds\right] = 0.$$

This implies that the sequences $\{u_t^n\}$, $\{V_t^n\}$, and $\{\sigma_t^n\}$ are Cauchy sequences in the appropriate spaces. More precisely, $\{u_t^n\}$ is a Cauchy sequence in the

space of progressively measurable processes u such that $\mathbb{E}[\int_0^T \|u_t\|^2 dt] < \infty$, $\{V_t^n\}$ is a Cauchy sequence in the space of progressively measurable processes V such that $\mathbb{E}[\int_0^T |V_t|^2 dt] < \infty$, and $\{\sigma_t^n\}$ is a Cauchy sequence in the space of progressively measurable processes σ such that $\mathbb{E}[\int_0^T \|\sigma_t\|^2 dt] < \infty$. Since these spaces are complete, there exist limit processes u_t, V_t, and σ_t such that $\lim_{n\to\infty} \mathbb{E}\left[\int_0^T \|u_t^n - u_t\|^2 dt\right] = 0$, $\lim_{n\to\infty} \mathbb{E}\left[\int_0^T |V_t^n - V_t|^2 dt\right] = 0$, $\lim_{n\to\infty} \mathbb{E}\left[\int_0^T \|\sigma_t^n - \sigma_t\|^2 dt\right] = 0$. Moreover, by passing to the limit in the iterative scheme, we can show that (u_t, V_t, σ_t) is a solution to the coupled FBSDE system (Forward Propagation)-(Backward Propagation).

Uniqueness of the Solution: Suppose there are two solutions (u_t, V_t, σ_t) and $(\tilde{u}_t, \tilde{V}_t, \tilde{\sigma}_t)$ to the coupled FBSDEs (Forward Propagation)-(Backward Propagation). Let $\Delta u_t = u_t - \tilde{u}_t$, $\Delta V_t = V_t - \tilde{V}_t$, and $\Delta\sigma_t = \sigma_t - \tilde{\sigma}_t$. Then, we have:

$$d(\Delta u_t) = [h_{\theta_t}(u_t, c_t) - h_{\theta_t}(\tilde{u}_t, c_t)]dt + [\sigma^h_{\theta_t}(u_t, c_t) - \sigma^h_{\theta_t}(\tilde{u}_t, c_t)]d\mathcal{Z}^h_t,$$
$$d(\Delta V_t) = -[f_{\theta_t}(c_t, u_t, V_t) - f_{\theta_t}(c_t, \tilde{u}_t, \tilde{V}_t) - \lambda(\mathcal{L}(u_t, \theta_t) - \mathcal{L}(\tilde{u}_t, \theta_t))]dt + \Delta\sigma_t d\mathcal{Z}_t.$$

Using the same arguments as in the derivation of the estimates for δu_t^n and δV_t^n, we can show that

$$\mathbb{E}[\|\Delta u_t\|^2] \leq L_h \int_0^t \mathbb{E}[\|\Delta u_s\|^2] ds,$$

and

$$\mathbb{E}[|\Delta V_t|^2] + \mathbb{E}\left[\int_t^T \|\Delta\sigma_s\|^2 ds\right] \leq C_V \int_t^T \mathbb{E}[|\Delta V_s|^2] ds + C_V \int_t^T \mathbb{E}[\|\Delta u_s\|^2] ds.$$

By Gronwall's inequality, we have $\mathbb{E}[\|\Delta u_t\|^2] = 0$ for all $t \in [0, T]$, which implies $\Delta u_t = 0$ almost surely for all $t \in [0, T]$. Then, from the inequality for ΔV_t, we have $\mathbb{E}[|\Delta V_t|^2] = 0$ and $\mathbb{E}[\int_t^T \|\Delta\sigma_s\|^2 ds] = 0$, which implies $\Delta V_t = 0$ almost surely for all $t \in [0, T]$ and $\Delta\sigma_t = 0$ almost surely for almost every $t \in [0, T]$. Thus, the solution to the coupled FBSDE system is unique.

Part 2: Existence of an Optimal Control θ_t^*. We now adapt the proof by using a combination of dynamic programming, weak convergence arguments, and a verification theorem.

Step 1: Define a Value Function Define the value function $V_t(c)$ as:

$$V_t(c) = \sup_{\theta \in \mathcal{A}} \mathbb{E}\left[\int_t^T \left(f_{\theta_s}(c_s, u_s, V_s) - \lambda \mathcal{L}(u_s, \theta_s)\right) ds \middle| \mathcal{F}_t\right],$$

where $\mathcal{F}_t$ represents the information available at time t, including the history of the consumption process c up to time t. This is the agent's optimization problem starting from time t, given the consumption process c.

Step 2: Apply Dynamic Programming Principle By the dynamic programming principle (see, e.g., Fleming and Rishel (2012)), we have:

$$V_t(c) = \sup_{\theta \in \mathcal{A}} \mathbb{E}\left[\int_t^{t+\Delta t} \left(f_{\theta_s}(c_s, u_s, V_s) - \lambda \mathcal{L}(u_s, \theta_s)\right) ds + V_{t+\Delta t}(c) \middle| \mathcal{F}_t\right],$$

for any $\Delta t > 0$ such that $t + \Delta t \leq T$. Note that $V_{t+\Delta t}(c)$ depends on the consumption process c from t to T.

Step 3: Derive the HJB Equation Deriving the HJB equation in terms of $V_t(c)$ is complex because it involves functional derivatives with respect to the consumption process. However, for practical purposes and numerical approximations, we can often work with a discretized version of the HJB equation or consider the HJB equation in terms of the state variable u_t and the history of consumption up to time t. Formally, we can write the HJB equation as:

$$\sup_{\theta \in \Theta} \left\{ \frac{\partial V}{\partial t} + \mathbb{E}\left[\frac{\partial V}{\partial c}\frac{dc}{dt} \middle| \mathcal{F}_t\right] + \frac{\partial V}{\partial u} h_\theta(u, c) + \frac{1}{2}\text{tr}\left(H_u(V)(\sigma_\theta^h(u, c))(\sigma_\theta^h(u, c))^T\right) + f_\theta(c, u, V) - \lambda \mathcal{L}(u, \theta) \right\} = 0,$$

with terminal condition $V_T(c) = 0$. Here, $\frac{\delta V}{\delta c}$ represents a functional derivative of the value function with respect to the consumption process.

Step 4: Verify Assumptions for HJB Equation We need to show that the value function is sufficiently smooth and satisfies certain growth conditions for the HJB equation to be well-defined. In practice, we often rely on numerical approximations or make simplifying assumptions about the structure of the problem to make it more tractable (see, e.g., Fleming and Rishel (2012), or Bardi et al. (1997) for a discussion of viscosity solutions and their regularity properties).

Step 5: Constructing a Maximizing Sequence and Using Weak Compactness Let $\{\theta^n\}$ be a maximizing sequence of controls in $\mathcal{A}$ such that

$$\lim_{n\to\infty} J(t,c;\theta^n) = V_t(c),$$

where $J(t,c;\theta)$ is the objective functional in (18) starting from time t under consumption process c and control θ. Since Θ is compact and the mapping $s \mapsto \theta_s^n$ is measurable for each n, by the Dunford-Pettis theorem and the Komlós theorem (see, e.g., Ekeland and Temam (1999)), there exists a subsequence (still denoted by $\{\theta^n\}$) and a measurable map $\theta^* : [0,T] \times \Omega \to \Theta$ such that θ^n converges to θ^* in a weak sense. Specifically, for any $t \in [0,T]$ and any integrable function g, we have

$$\lim_{n\to\infty} \mathbb{E}\left[\int_t^T g(s)\theta_s^n ds\right] = \mathbb{E}\left[\int_t^T g(s)\theta_s^* ds\right].$$

Since the space of admissible controls $\mathcal{A}$ consists of progressively measurable processes with values in Θ, and Θ is convex, the weak limit θ^* is also an admissible control.[28]

Step 6: Showing that the Weak Limit is the Optimal Control Let (u^n, V^n) be the solution to the FBSDE system corresponding to θ^n and a given consumption process c, and let (u^*, V^*) be the solution corresponding to θ^* and the same consumption process c. We want to show that $J(t,c;\theta^*) = V_t(c)$.[29] By the stability properties of

[28] The specific type of weak convergence implied by the Dunford-Pettis and Komlós theorems in this context is weak convergence in $L^1([t,T] \times \Omega; \Theta)$. This means that for any function $g \in L^\infty([t,T] \times \Omega)$, the above convergence holds.

[29] To show the convergence of (u^n, V^n) to (u^*, V^*), we need a stability result for FBSDEs under weak convergence of controls. This is a non-trivial result and typically requires additional assumptions on the coefficients of the FBSDE. For example, if the coefficients h, σ^h, f, and $\mathcal{L}$ are Lipschitz continuous

FBSDEs (as mentioned in the remark above), and using the Lipschitz continuity of h, f, $\mathcal{L}$ and the uniform boundedness of σ^h, we can show that (u^n, V^n) converges to (u^*, V^*) in an appropriate sense as $n \to \infty$. Specifically, we can show that $\mathbb{E}[\sup_{t \le s \le T} \|u_s^n - u_s^*\|^2] \to 0$ and $\mathbb{E}[\int_t^T |V_s^n - V_s^*|^2 ds] \to 0$. Now, consider the difference $|J(t,c;\theta^n) - J(t,c;\theta^*)|$. Using the Lipschitz continuity of f_θ and $\mathcal{L}$ in θ, and the weak convergence of θ^n to θ^* established in Step 5, we have

$$\lim_{n\to\infty} \mathbb{E}\left[\int_t^T f_{\theta_s^n}(c_s, u_s^n, V_s^n) ds\right] = \mathbb{E}\left[\int_t^T f_{\theta_s^*}(c_s, u_s^*, V_s^*) ds\right],$$
$$\lim_{n\to\infty} \mathbb{E}\left[\int_t^T \mathcal{L}(u_s^n, \theta_s^n) ds\right] = \mathbb{E}\left[\int_t^T \mathcal{L}(u_s^*, \theta_s^*) ds\right].$$

Combining these with the convergence of (u^n, V^n) to (u^*, V^*), we can show that

$$\begin{aligned} |J(t,c;\theta^n) - J(t,c;\theta^*)| &\le \mathbb{E}\left[\int_t^T |f_{\theta_s^n}(c_s, u_s^n, V_s^n) - f_{\theta_s^*}(c_s, u_s^*, V_s^*)| ds\right] \\ &\quad + \lambda \mathbb{E}\left[\int_t^T |\mathcal{L}(u_s^n, \theta_s^n) - \mathcal{L}(u_s^*, \theta_s^*)| ds\right] \le \epsilon_n, \end{aligned}$$

where $\epsilon_n \to 0$ as $n \to \infty$.

Detailed Derivation of the Inequality: We can bound the first term on the right-hand side as follows:

$$\begin{aligned} &\mathbb{E}\left[\int_t^T |f_{\theta_s^n}(c_s, u_s^n, V_s^n) - f_{\theta_s^*}(c_s, u_s^*, V_s^*)| ds\right] \\ &\le \mathbb{E}\left[\int_t^T |f_{\theta_s^n}(c_s, u_s^n, V_s^n) - f_{\theta_s^n}(c_s, u_s^*, V_s^*)| ds\right] \\ &+ \mathbb{E}\left[\int_t^T |f_{\theta_s^n}(c_s, u_s^*, V_s^*) - f_{\theta_s^*}(c_s, u_s^*, V_s^*)| ds\right] \\ &\le L_f \mathbb{E}\left[\int_t^T (\|u_s^n - u_s^*\| + |V_s^n - V_s^*|) ds\right] + L_{f,\theta} \mathbb{E}\left[\int_t^T \|\theta_s^n - \theta_s^*\| ds\right] \end{aligned}$$

---

in θ, and if σ_θ^h is uniformly bounded, then we can show that the solution map $\theta \mapsto (u, V)$ is continuous from $L^p([t,T] \times \Omega)$ (for some $p \ge 1$) under the weak topology to $\mathcal{S}^2 \times \mathcal{H}^2$ under the strong topology. This result can be found in, (e.g., Yong and Zhou (2012)) or similar references on FBSDEs.

$$\leq L_f \left(\mathbb{E}\left[\int_t^T \|u_s^n - u_s^*\|^2 ds \right] \right)^{1/2} \sqrt{T-t} + L_f \left(\mathbb{E}\left[\int_t^T |V_s^n - V_s^*|^2 ds \right] \right)^{1/2} \sqrt{T-t}$$
$$+ L_{f,\theta} \mathbb{E}\left[\int_t^T \|\theta_s^n - \theta_s^*\| ds \right].$$

where L_f is the Lipschitz constant of f in (u, V) and $L_{f,\theta}$ is the Lipschitz constant of f in θ. The first two terms go to 0 as $n \to \infty$ due to the strong convergence of (u^n, V^n) to (u^*, V^*), and the last term goes to 0 due to the weak convergence of θ^n to θ^*. To elaborate, the weak convergence of θ^n to θ^* in L^1 implies that for any bounded measurable function g, we have

$$\lim_{n\to\infty} \mathbb{E}\left[\int_t^T g(s)\theta_s^n ds \right] = \mathbb{E}\left[\int_t^T g(s)\theta_s^* ds \right].$$

Since Θ is compact, it is also bounded. Let M be a bound such that $\|\theta\| \leq M$ for all $\theta \in \Theta$. Then by the Cauchy-Schwarz inequality, we have

$$\mathbb{E}\left[\int_t^T \|\theta_s^n - \theta_s^*\| ds \right] \leq \sqrt{T-t} \left(\mathbb{E}\left[\int_t^T \|\theta_s^n - \theta_s^*\|^2 ds \right] \right)^{1/2}.$$

Note that $\|\theta_s^n - \theta_s^*\|^2 \leq (2M)^2 = 4M^2$. Thus by dominated convergence theorem, we have

$$\lim_{n\to\infty} \mathbb{E}\left[\int_t^T \|\theta_s^n - \theta_s^*\|^2 ds \right] = 0, \quad \text{thus,} \quad \lim_{n\to\infty} \mathbb{E}\left[\int_t^T \|\theta_s^n - \theta_s^*\| ds \right] = 0.$$

Similarly, we can bound the second term using the Lipschitz continuity of $\mathcal{L}$ in u and θ. Since $\{\theta^n\}$ is a maximizing sequence, we have $\lim_{n\to\infty} J(t, c; \theta^n) = V_t(c)$. Taking the limit as $n \to \infty$ in the inequality above, we get

$$|V_t(c) - J(t, c; \theta^*)| \leq 0,$$

which implies $V_t(c) = J(t, c; \theta^*)$. Thus, θ^* is an optimal control.

Step 7: Applying a Verification Theorem. To formally establish the optimality of the control θ^*, we employ a verification theorem from stochastic optimal control

theory (e.g., Yong and Zhou (2012), Fleming and Rishel (2012)). This theorem asserts that if a sufficiently regular candidate value function satisfies the Hamilton-Jacobi-Bellman (HJB) equation, it is indeed the value function, and any control achieving the supremum in the HJB equation is optimal. Under Assumption 7, we consider the HJB equation associated with our problem. Assuming sufficient regularity of the value function $V(t,c)$, the HJB equation is:

$$-\frac{\partial V}{\partial t}(t,c) = \sup_{\theta\in\Theta}\left\{\mathcal{A}^{\theta}V(t,c) + f_{\theta}(c,u^*(t,c),V(t,c)) - \lambda\mathcal{L}(u^*(t,c),\theta)\right\},$$

with the terminal condition $V(T,c)=0$. Here, $\mathcal{A}^{\theta}$ denotes the infinitesimal generator of the state process u_t under control θ, and $u^*(t,c)$ is the state process under the optimal strategy. Consider the value function defined by:

$$V(t,c) = \sup_{\theta\in\mathcal{A}} J(t,c;\theta).$$

If this value function is sufficiently regular and satisfies the HJB equation, it can verify the optimality of θ^*. For an arbitrary admissible control $\theta\in\mathcal{A}$, let $(u_t^{\theta},V_t^{\theta})$ be the corresponding FBSDE solution. Applying Itô's lemma to $V(t,c_t)$ along the state process u_t^{θ} and using the HJB equation's supremum property yields:

$$-dV(t,c_t) \geq \left\{\mathcal{A}^{\theta_t}V(t,c_t) + f_{\theta_t}(c_t,u_t^{\theta},V_t^{\theta}) - \lambda\mathcal{L}(u_t^{\theta},\theta_t)\right\}dt - d\mathcal{Z}_t,$$

where $\mathcal{Z}_t$ is a martingale. Integrating from t to T, taking expectations, and using $V(T,c_T)=0$:

$$V(t,c_t) \geq \mathbb{E}_t\left[\int_t^T \left(f_{\theta_s}(c_s,u_s^{\theta},V_s^{\theta}) - \lambda\mathcal{L}(u_s^{\theta},\theta_s)\right)ds\right] = J(t,c;\theta).$$

This confirms $V(t,c)$ as an upper bound for the value functional. For the control θ^* from Step 5, which achieves the supremum in the HJB equation almost everywhere, the inequality becomes an equality:

$$-dV(t,c_t) = \left\{\mathcal{A}^{\theta_t^*}V(t,c_t) + f_{\theta_t^*}(c_t,u_t^{\theta^*},V_t^{\theta^*}) - \lambda\mathcal{L}(u_t^{\theta^*},\theta_t^*)\right\}dt - d\mathcal{Z}_t^*.$$

Integrating and taking expectations:

$$V(t,c_t) = \mathbb{E}_t\left[\int_t^T \left(f_{\theta_s^*}(c_s, u_s^{\theta^*}, V_s^{\theta^*}) - \lambda\mathcal{L}(u_s^{\theta^*}, \theta_s^*)\right) ds\right] = J(t,c;\theta^*).$$

Thus, the value function $V(t,c)$ is achieved by θ^*, making it an optimal control. Under the stated assumptions and the existence of a suitable HJB solution, the verification theorem confirms the optimality of the constructed control θ^*.

Q.E.D.

REFERENCES

Abel, Andrew B (1990): "Asset prices under habit formation and catching up with the Joneses," *American Economic Review*, 80 (2), 38–42. [2]

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2022): "Income and wealth distribution in macroeconomics: A continuous-time approach," *The review of economic studies*, 89 (1), 45–86. [75]

Ahuja, Saran (2016): "Wellposedness of mean field games with common noise under a weak monotonicity condition," *SIAM Journal on Control and Optimization*, 54 (1), 30–48. [9]

Allais, Maurice (1953): "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine," *Econometrica: journal of the Econometric Society*, 503–546. [31]

Alvarez, Fernando, Francesco Lippi, and Panagiotis Souganidis (2023): "Price setting with strategic complementarities as a mean field game," *Econometrica*, 91 (6), 2005–2039. [75]

Anthony, Martin, Peter L Bartlett, Peter L Bartlett, et al. (1999): *Neural network learning: Theoretical foundations*, vol. 9, cambridge university press Cambridge. [16]

Arrow, Kenneth J. (1963): "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review*, 53 (5), 941–973. [2]

Baley, Isaac and Laura Veldkamp (2023): "Chapter 23 - Bayesian learning." in *Handbook of Economic Expectations*, ed. by Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw, Academic Press, 717–748. [7]

Bardi, Martino, Italo Capuzzo Dolcetta, et al. (1997): *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, vol. 12, Springer. [41]

Bechara, Antoine (2004): "The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage," *Brain and cognition*, 55 (1), 30–40. [63]

Bhui, Rahul, Lucy Lai, and Samuel J Gershman (2021): "Resource-rational decision making," *Current Opinion in Behavioral Sciences*, 41, 15–21. [2, 31]

Bolton, Patrick, Hui Chen, and Neng Wang (2011): "A unified theory of Tobin's q, corporate investment, financing, and risk management," *The journal of Finance*, 66 (5), 1545–1578. [12]

Borovička, Jaroslav (2020): "Survival and long-run dynamics with heterogeneous beliefs under recursive preferences," *Journal of Political Economy*, 128 (1), 206–251. [7]

Botvinick, Matthew M, Todd S Braver, Deanna M Barch, Cameron S Carter, and Jonathan D Cohen (2001): "Conflict monitoring and cognitive control." *Psychological review*, 108 (3), 624. [5, 61]

Brocas, Isabelle and Juan D Carrillo (2008): "The brain as a hierarchical organization," *American Economic Review*, 98 (4), 1312–46. [3]

Camerer, Colin, George Loewenstein, and Drazen Prelec (2005): "Neuroeconomics: How neuroscience can inform economics," *Journal of economic Literature*, 43 (1), 9–64. [2, 3, 4, 31]

Camerer, Colin F (2004): *Advances in Behavioral Economics*, Princeton University Press. [2]

Carmona, René and François Delarue (2015): "Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics," *The Annals of Probability*, 2647–2700. [9]

Chen, Zengjing and Larry Epstein (2002): "Ambiguity, risk, and asset returns in continuous time," *Econometrica*, 70 (4), 1403–1443. [7, 8]

Cybenko, George (1989): "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, 2 (4), 303–314. [68]

Debreu, Gerard et al. (1954): "Representation of a preference ordering by a numerical function," *Decision processes*, 3, 159–165. [1]

Duffie, Darrell and Larry G Epstein (1992): "Stochastic differential utility," *Econometrica: Journal of the Econometric Society*, 353–394. [6, 14, 15, 18, 26, 32, 34]

Duffie, Darrell and Costis Skiadas (1994): "Continuous-time security pricing: A utility gradient approach," *Journal of Mathematical Economics*, 23 (2), 107–131. [18]

Eisner, Robert (1963): "Investment: Fact and fancy," *The American Economic Review*, 53 (2), 237–246. [62]

Ekeland, Ivar and Roger Temam (1999): *Convex analysis and variational problems*, SIAM. [41]

Ellsberg, Daniel (1961): "Risk, ambiguity, and the Savage axioms," *Quarterly journal of economics*, 643–669. [2, 31]

Epstein, Larry and Tan Wang (1994): "Intertemporal Asset Pricing Under Knightian Uncertainty," *Econometrica*, 62 (2), 283–322. [7, 8]

Epstein, Larry G and Shaolin Ji (2013): "Ambiguous volatility and asset pricing in continuous time," *The Review of Financial Studies*, 26 (7), 1740–1786. [7, 8]

Epstein, Larry G. and Stanley E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57 (4), 937–969. [6, 14, 18]

Fleming, Wendell H and Raymond W Rishel (2012): *Deterministic and stochastic optimal control*, vol. 1, Springer Science & Business Media. [40, 41, 44]

Frederick, Shane, George Loewenstein, and Ted O'donoghue (2002): "Time discounting and time preference: A critical review," *Journal of economic literature*, 40 (2), 351–401. [2]

Gabaix, Xavier (2019): "Behavioral inattention," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2, 261–343. [2, 66]

Gabaix, Xavier and Thomas Graeber (2024): "The complexity of economic decisions," Tech. rep., National Bureau of Economic Research. [3, 66]

Gershman, Samuel J, Eric J Horvitz, and Joshua B Tenenbaum (2015): "Computational rationality: A converging paradigm for intelligence in brains, minds, and machines," *Science*, 349 (6245), 273–278. [2, 31]

Gigerenzer, Gerd and Wolfgang Gaissmaier (2011): *Heuristics: The foundations of adaptive behavior*, Oxford University Press. [62]

Gilboa, Itzhak and David Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of mathematical economics*, 18 (2), 141–153. [7, 8]

Glimcher, Paul W (2013): *Neuroeconomics: Decision making and the brain*, Academic Press. [67]

Glorot, Xavier and Yoshua Bengio (2010): "Understanding the difficulty of training deep feedforward neural networks," *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 249–256. [68]

Goodfellow, Ian (2016): "Deep learning," . [5, 63, 68, 69]

Gross, James J (1998): "Antecedent-and response-focused emotion regulation: divergent consequences for experience, expression, and physiology." *Journal of personality and social psychology*, 74 (1), 224. [5, 61]

——— (2002): "Emotion regulation: Affective, cognitive, and social consequences," *Psychophysiology*, 39 (3), 281–291. [62]

Haber, Eldad and Lars Ruthotto (2017): "Stable architectures for deep neural networks," *Inverse problems*, 34 (1), 014004. [3, 15]

Hansen, Lars Peter and Thomas J Sargent (2001): "Robust control and model uncertainty," *American Economic Review*, 91 (2), 60–66. [7, 8, 16]

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (2016): "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778. [3, 15, 67]

Hornik, Kurt, Maxwell Stinchcombe, and Halbert White (1989): "Multilayer feedforward networks are universal approximators," *Neural networks*, 2 (5), 359–366. [68]

Ikeda, Nobuyuki and Shinzo Watanabe (2014): *Stochastic differential equations and diffusion processes*, Elsevier. [56]

Kahneman, Daniel (2011): "Thinking, fast and slow," *Farrar, Straus and Giroux*. [2, 3, 4, 5, 31, 61, 67]

Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler (1991): "Anomalies: The endowment effect, loss aversion, and status quo bias," *Journal of Economic perspectives*, 5 (1), 193–206. [2]

Kahneman, DANIEL and Amos Tversky (1979): "Prospect theory: An analysis of decision under risk," *Econometrica*, 47 (2), 363–391. [2, 31]

Koopmans, Tjalling C (1960): "Stationary ordinal utility and impatience," *Econometrica: Journal of the Econometric Society*, 287–309. [7]

Kraft, Holger and Frank Thomas Seifried (2014): "Stochastic differential utility as the continuous-time limit of recursive utility," *Journal of Economic Theory*, 151, 528–550. [18]

Kreps, David M and Evan L Porteus (1978): "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica: journal of the Econometric Society*, 185–200. [6, 14, 18]

Laibson, David (1997): "Golden eggs and hyperbolic discounting," *The Quarterly Journal of Economics*, 112 (2), 443–478. [2]

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015): "Deep learning," *nature*, 521 (7553), 436–444. [3, 67]

Li, Qianxiao, Long Chen, Cheng Tai, and E Weinan (2018): "Maximum principle based algorithms for deep learning," *Journal of Machine Learning Research*, 18 (165), 1–29. [3, 15, 21]

Li, Qianxiao, Ting Lin, and Zuowei Shen (2022): "Deep learning via dynamical systems: An approximation perspective," *Journal of the European Mathematical Society*. [3, 14, 15]

Lieder, Falk and Thomas L Griffiths (2020): "Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources," *Behavioral and brain sciences*, 43, e1. [2, 31]

Loewenstein, George (2000): "Emotions in economic theory and economic behavior," *American Economic Review*, 90 (2), 426–432. [61]

Lu, LuShin, Yeonjong Shin, Yanhui Su, and George Em Karniadakis (2020): "Dying ReLU and initialization: Theory and numerical examples," *Communications in Computational Physics*, 28 (5), 1671–1706. [70]

Lucas, Robert E and Thomas J Sargent (1978): "After keynesian macroeconomics," *After the Phillips curve: Persistence of high inflation and high unemployment*, 19, 49–72. [2]

Lucas Jr, Robert E (1967): "Adjustment costs and the theory of supply," *Journal of political economy*, 75 (4, Part 1), 321–334. [4, 62]

——— (1976): "Econometric policy evaluation: A critique," in *Carnegie-Rochester conference series on public policy*, North-Holland, vol. 1, 19–46. [2]

——— (1978): "Asset prices in an exchange economy," *Econometrica: journal of the Econometric Society*, 1429–1445. [2]

Ma, Jin and Jiongmin Yong (2007): *Forward-Backward Stochastic Differential Equations and their Applications*, Springer. [74]

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini (2006): "Ambiguity aversion, robustness, and the variational representation of preferences," *Econometrica*, 74 (6), 1447–1498. [7, 8, 16]

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt (2023): "Rational inattention: A review," *Journal of Economic Literature*, 61 (1), 226–273. [6]

Malmendier, Ulrike and Stefan Nagel (2011): "Depression babies: do macroeconomic experiences affect risk taking?" *The quarterly journal of economics*, 126 (1), 373–416. [2]

Mayer, John D, Richard D Roberts, and Sigal G Barsade (2008): "Human abilities: Emotional intelligence," *Annu. Rev. Psychol.*, 59 (1), 507–536. [5, 61]

McFadden, Daniel (2001): "Economic choices," *American economic review*, 91 (3), 351–378. [6]

Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. (2015): "Human-level control through deep reinforcement learning," *nature*, 518 (7540), 529–533. [14, 72]

O'donoghue, Ted and Matthew Rabin (1999): "Doing it now or later," *American economic review*, 89 (1), 103–124. [2]

Ogaki, Masao and Qiang Zhang (2001): "Decreasing relative risk aversion and tests of risk sharing," *Econometrica*, 69 (2), 515–526. [13]

Pardoux, Etienne and Shige Peng (1990): "Adapted solution of a backward stochastic differential equation," *Systems & control letters*, 14 (1), 55–61. [24, 34, 57]

Peng, Shige (1997): "Backward SDE and related g-expectation," *Pitman research notes in mathematics series*, 141–160. [8]

——— (2007): "G-expectation, G-Brownian motion and related stochastic calculus of Itô type," in *Stochastic Analysis and Applications: The Abel Symposium 2005*, Springer, 541–567. [8]

Peng, Shige and Zhen Wu (1999): "Fully coupled forward-backward stochastic differential equations and applications to optimal control," *SIAM Journal on Control and Optimization*, 37 (3), 825–843. [9, 55]

Post, Thierry, Martijn J Van den Assem, Guido Baltussen, and Richard H Thaler (2008): "Deal or no deal? decision making under risk in a large-payoff game show," *American Economic Review*, 98 (1), 38–71. [13]

Prescott, Edward C. and Rajnish Mehra (1980): "Recursive Competitive Equilibrium: The Case of Homogeneous Households," *Econometrica*, 48 (6), 1365–1379. [2]

Qi, Qian (2025a): "Adjustment Theory," *SSRN Working paper*. [4, 5, 31]

——— (2025b): "Dynamic Time Inconsistency," *SSRN Working paper*. [5, 31]

——— (2025c): "Endogenous Ambiguity Aversion," *SSRN Working paper*. [4, 31]

——— (2025d): "Universal Preference Approximation," *SSRN Working paper*. [3, 14, 31]

Rabin, Matthew (1998): "Psychology and economics," *Journal of economic literature*, 36 (1), 11–46. [2]

——— (2000): "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 68 (5), 1281–1292. [13]

Rescorla, Robert A and Allan R Wagner (1972): "A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement," *Classical Conditioning II: Current Research and Theory*, 3, 64–99. [66]

Salovey, Peter and John D Mayer (1990): "Emotional intelligence," *Imagination, cognition and personality*, 9 (3), 185–211. [5, 61]

Sargent, Thomas J (1993): *Bounded rationality in macroeconomics*, Clarendon Press. [62, 66]

——— (2024a): "Critique and consequence," *Journal of Monetary Economics*, 141, 2–13. [2]

——— (2024b): "Macroeconomics After Lucas," *Working Paper*. [2]

Schmeidler, David (1989): "Subjective probability and expected utility without additivity," *Econometrica: Journal of the Econometric Society*, 571–587. [7, 8]

Shenhav, Amitai, Matthew M Botvinick, and Jonathan D Cohen (2013): "The expected value of control: an integrative theory of anterior cingulate cortex function," *Neuron*, 79 (2), 217–240. [30]

Shenhav, Amitai, Jonathan D Cohen, and Matthew M Botvinick (2016): "Dorsal anterior cingulate cortex and the value of control," *Nature neuroscience*, 19 (10), 1286–1291. [30]

Simon, Herbert A (1955): "A behavioral model of rational choice," *The quarterly journal of economics*, 69 (1), 99–118. [7, 61, 62]

——— (1972): "Theories of bounded rationality," *Decision and Organization: A Volume in Honor of Jacob Marschak/North Holland*. [61]

Sims, Christopher A (2003): "Implications of rational inattention," *Journal of monetary Economics*, 50 (3), 665–690. [4, 6, 16, 62, 63]

Skiadas, Costis (2003): "Robust control and recursive utility," *Finance and Stochastics*, 7, 475–489. [6, 8, 16]

Srivastava, Rupesh K, Klaus Greff, and Jürgen Schmidhuber (2015): "Training very deep networks," *Advances in neural information processing systems*, 28. [3]

Strotz, Robert H (1955): "Myopia and inconsistency in dynamic utility maximization," *The review of economic studies*, 23 (3), 165–180. [67]

Strzalecki, Tomasz (2011): "Axiomatic foundations of multiplier preferences," *Econometrica*, 79 (1), 47–73. [7, 8, 16]

Sutton, Richard S and Andrew G Barto (2018): *Reinforcement learning: An introduction*, MIT press. [72]

Swanson, Eric T (2012): "Risk aversion and the labor margin in dynamic equilibrium models," *American Economic Review*, 102 (4), 1663–1691. [13]

Tesfatsion, Leigh (2002): "Agent-based computational economics: growing economies from the bottom up," *Artificial life*, 8 (1), 55–82. [62, 67]

Uzawa, Hirofumi (1968): *Time Preference, the Consumption Function, and Optimum Asset Holdings. In: JN Wolfe, ed. Value, Capital, and Growth: Papers in Honor of Sir John Hicks. Edinburgh. Scotland: University of Edinburgh Press*, Edinburgh, Scotland: University of Edinburgh Press. [7]

Verrecchia, Robert E (1982): "Information acquisition in a noisy rational expectations economy," *Econometrica: Journal of the Econometric Society*, 1415–1430. [63]

Watkins, Christopher JCH and Peter Dayan (1992): "Q-learning," *Machine learning*, 8, 279–292. [72]

Weinan, E (2017): "A proposal on machine learning via dynamical systems," *Communications in Mathematics and Statistics*, 1 (5), 1–11. [3, 15]

Weinan, E, Jiequn Han, and Qianxiao Li (2019): "A mean-field optimal control formulation of deep learning," *Research in the Mathematical Sciences*, 6 (1), 1–41. [3, 9, 15, 21, 22]

Yong, Jiongmin and Xun Yu Zhou (2012): *Stochastic controls: Hamiltonian systems and HJB equations*, vol. 43, Springer Science & Business Media. [26, 42, 44, 73, 74]

Zhang, Jianfeng (2004): "A numerical scheme for BSDEs," *The annals of applied probability*, 14 (1), 459–488. [24]

SUPPLEMENT TO "NEURAL FOUNDATIONS"

2. OMITTED PROOFS

PROOF OF PROPOSITION 2: We prove the proposition in several steps:

Step 1: Existence and Uniqueness. Let $u_t^0 = \phi(c_0)$, $u_t^{n+1} = \phi(c_0) + \int_0^t h_{\theta_s}(u_s^n, c_s)ds + \int_0^t \sigma_{\theta_s}^h(u_s^n, c_s)d\mathcal{Z}_s^h$. Fix $p \geq 2$. Let $y_n(t) = \mathbb{E}[\sup_{0\leq r\leq t}\|u_r^n - u_r^{n-1}\|^p]$. By Lipschitz continuity of h, σ, and Jensen's and BDG inequalities, for $n \geq 1$,

$$
\begin{aligned}
y_{n+1}(t) &\leq 2^{p-1}\mathbb{E}\left[\sup_{0\leq r\leq t}\left\|\int_0^r h_{\theta_s}(u_s^n, c_s) - h_{\theta_s}(u_s^{n-1}, c_s)ds\right\|^p\right] \\
&\quad + 2^{p-1}\mathbb{E}\left[\sup_{0\leq r\leq t}\left\|\int_0^r \sigma_{\theta_s}^h(u_s^n, c_s) - \sigma_{\theta_s}^h(u_s^{n-1}, c_s)d\mathcal{Z}_s^h\right\|^p\right] \\
&\leq 2^{p-1}T^{p-1}\mathbb{E}\left[\int_0^t \|h_{\theta_s}(u_s^n, c_s) - h_{\theta_s}(u_s^{n-1}, c_s)\|^p ds\right] \\
&\quad + 2^{p-1}C_{BDG}(p)\mathbb{E}\left[\left(\int_0^t \|\sigma_{\theta_s}^h(u_s^n, c_s) - \sigma_{\theta_s}^h(u_s^{n-1}, c_s)\|^2 ds\right)^{p/2}\right] \\
&\leq 2^{p-1}T^{p-1}K^p\int_0^t \mathbb{E}\left[\|u_s^n - u_s^{n-1}\|^p\right] ds \\
&\quad + 2^{p-1}C_{BDG}(p)T^{p/2-1}\mathbb{E}\left[\int_0^t \|\sigma_{\theta_s}^h(u_s^n, c_s) - \sigma_{\theta_s}^h(u_s^{n-1}, c_s)\|^p ds\right] \\
&\leq 2^{p-1}T^{p-1}K^p\int_0^t \mathbb{E}\left[\|u_s^n - u_s^{n-1}\|^p\right] ds \\
&\quad + 2^{p-1}C_{BDG}(p)T^{p/2-1}K^p\int_0^t \mathbb{E}\left[\|u_s^n - u_s^{n-1}\|^p\right] ds \\
&\leq 2^{p-1}T^{p-1}K^p\int_0^t \mathbb{E}\left[\sup_{0\leq r\leq s}\|u_r^n - u_r^{n-1}\|^p\right] ds \\
&\quad + 2^{p-1}C_{BDG}(p)T^{p/2-1}K^p\int_0^t \mathbb{E}\left[\sup_{0\leq r\leq s}\|u_r^n - u_r^{n-1}\|^p\right] ds \\
&\leq C_1(T,p)\int_0^t y_n(s)ds,
\end{aligned}
$$

where $C_1(T,p) = 2^{p-1}(T^{p-1} + C_{BDG}(p)T^{p/2-1})K^p$.

Justification of Hölder's Inequality: Applying Hölder's inequality with exponents $p/2$ and q where $1/q + 2/p = 1$, we get:

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t \|\sigma^h_{\theta_s}(u^n_s, c_s) - \sigma^h_{\theta_s}(u^{n-1}_s, c_s)\|^2 ds\right)^{p/2}\right] \\
&\leq \mathbb{E}\left[\left(\int_0^t 1^q ds\right)^{p/(2q)} \left(\int_0^t \|\sigma^h_{\theta_s}(u^n_s, c_s) - \sigma^h_{\theta_s}(u^{n-1}_s, c_s)\|^p ds\right)\right] \\
&= T^{p/(2q)}\mathbb{E}\left[\int_0^t \|\sigma^h_{\theta_s}(u^n_s, c_s) - \sigma^h_{\theta_s}(u^{n-1}_s, c_s)\|^p ds\right] \\
&= T^{p/2-1}\mathbb{E}\left[\int_0^t \|\sigma^h_{\theta_s}(u^n_s, c_s) - \sigma^h_{\theta_s}(u^{n-1}_s, c_s)\|^p ds\right]
\end{aligned}$$

Then, using the Lipschitz condition on σ^h:

$$\leq T^{p/2-1}K^p \int_0^t \mathbb{E}\left[\|u^n_s - u^{n-1}_s\|^p\right] ds.$$

Iteration and $y_1(t)$ bound: Iterating, we have

$$\begin{aligned}
y_{n+1}(t) &\leq \int_0^t C_1(T,p) y_n(s) ds \\
&\leq \int_0^t C_1(T,p) \int_0^s C_1(T,p) y_{n-1}(r) dr ds \\
&\leq \cdots \\
&\leq \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} C_1(T,p)^n y_1(s_n) ds_n \cdots ds_2 ds_1 \\
&\leq M(T)\frac{(C_1(T,p)t)^n}{n!},
\end{aligned}$$

where the last inequality uses the bound for $y_1(t)$ derived below. By linear growth of h and σ^h, i.e., $\|h_{\theta_s}(u,c)\| \leq K_1(1 + \|u\| + \|c\|)$ and $\|\sigma^h_{\theta_s}(u,c)\| \leq K_2(1 + \|u\| + \|c\|)$,

$$y_1(t) \leq 2^{p-1}T^{p-1}\mathbb{E}\left[\int_0^t \|h_{\theta_s}(u^0_s, c_s)\|^p ds\right] + 2^{p-1}C_{BDG}(p)\mathbb{E}\left[\left(\int_0^t \|\sigma^h_{\theta_s}(u^0_s, c_s)\|^2 ds\right)^{p/2}\right]$$

$$\leq 2^{p-1}T^{p-1}K_1^p \int_0^t \mathbb{E}\left[(1+\|u_s^0\|+\|c_s\|)^p\right] ds$$

$$+ 2^{p-1}C_{BDG}(p)T^{p/2-1}K_2^p \int_0^t \mathbb{E}\left[(1+\|u_s^0\|+\|c_s\|)^p\right] ds$$

$$\leq 2^{p-1}3^{p-1}(T^{p-1}K_1^p + C_{BDG}(p)T^{p/2-1}K_2^p)\int_0^t \mathbb{E}\left[1+\|\phi(c_0)\|^p+\|c_s\|^p\right] ds$$

$$\leq C_2(T,p)T(1+\mathbb{E}[\|\phi(c_0)\|^p] + \sup_{0\leq s\leq T}\mathbb{E}[\|c_s\|^p]) =: M(T),$$

where $C_2(T,p) = 2^{p-1}3^{p-1}(T^{p-1}K_1^p + C_{BDG}(p)T^{p/2-1}K_2^p)$. Thus, $\mathbb{E}[\sup_{0\leq t\leq T}\|u_t^{n+1} - u_t^n\|^p] \leq \frac{(C_1(T,p)T)^n}{n!}M(T)$, so $\{u^n\}$ converges in $\mathcal{S}^p$ to a unique strong solution u_t.

Step 2: Continuity in c_0. Let $c_0^n \to c_0$, $u_0^n = \phi(c_0^n)$, $u_0 = \phi(c_0)$. By similar arguments,

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\|u_t^n - u_t\|^p\right]$$

$$\leq 2^{p-1}\mathbb{E}\left[\|\phi(c_0^n)-\phi(c_0)\|^p\right] + 2^{p-1}T^{p-1}\mathbb{E}\left[\int_0^T \|h_{\theta_s}(u_s^n,c_s) - h_{\theta_s}(u_s,c_s)\|^p ds\right]$$

$$+ 2^{p-1}C_{BDG}(p)T^{p/2-1}\mathbb{E}\left[\int_0^T \|\sigma^h_{\theta_s}(u_s^n,c_s) - \sigma^h_{\theta_s}(u_s,c_s)\|^p ds\right]$$

$$\leq 2^{p-1}\mathbb{E}\left[\|\phi(c_0^n)-\phi(c_0)\|^p\right] + 2^{p-1}T^{p-1}K^p\int_0^T \mathbb{E}\left[\|u_s^n - u_s\|^p\right] ds$$

$$+ 2^{p-1}C_{BDG}(p)T^{p/2-1}K^p\int_0^T \mathbb{E}\left[\|u_s^n - u_s\|^p\right] ds$$

$$\leq 2^{p-1}\mathbb{E}\left[\|\phi(c_0^n)-\phi(c_0)\|^p\right] + C_1(T,p)\int_0^T \mathbb{E}\left[\sup_{0\leq r\leq s}\|u_r^n - u_r\|^p\right] ds.$$

By Gronwall, $\mathbb{E}[\sup_{0\leq t\leq T}\|u_t^n - u_t\|^p] \leq 2^{p-1}\mathbb{E}[\|\phi(c_0^n)-\phi(c_0)\|^p]e^{C_1(T,p)T}$. Since ϕ is continuous, $\Phi_T(c_0^n;\theta) \to \Phi_T(c_0;\theta)$.

Step 3: Continuity in θ. Let $\theta^n \to \theta$ in $L^\infty([0,T];\Theta)$. Let u_t^n, u_t be solutions with θ^n, θ. Assume h is Lipschitz in θ, i.e., $\|h_{\theta^n}(u,c) - h_\theta(u,c)\| \leq L_h\|\theta^n - \theta\|$ for some constant L_h.

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\|u_t^n - u_t\|^p\right]$$

$$
\begin{aligned}
&\leq 2^{p-1}T^{p-1}\mathbb{E}\left[\int_0^T \|h_{\theta_s^n}(u_s^n,c_s)-h_{\theta_s}(u_s^n,c_s)\|^p+\|h_{\theta_s}(u_s^n,c_s)-h_{\theta_s}(u_s,c_s)\|^p ds\right]\\
&+2^{p-1}C_{BDG}(p)T^{p/2-1}\mathbb{E}\left[\int_0^T \|\sigma_{\theta_s}^h(u_s^n,c_s)-\sigma_{\theta_s}^h(u_s,c_s)\|^p ds\right]\\
&\leq 2^{p-1}T^{p-1}L_h^p\mathbb{E}\left[\int_0^T \|\theta_s^n-\theta_s\|^p ds\right]+2^{p-1}T^{p-1}K^p\int_0^T\mathbb{E}\left[\sup_{0\leq\tau\leq s}\|u_\tau^n-u_\tau\|^p\right]ds\\
&+2^{p-1}C_{BDG}(p)T^{p/2-1}K^p\int_0^T\mathbb{E}\left[\sup_{0\leq\tau\leq s}\|u_\tau^n-u_\tau\|^p\right]ds\\
&\leq 2^{p-1}T^pL_h^p\mathbb{E}\left[\sup_{0\leq s\leq T}\|\theta_s^n-\theta_s\|^p\right]+C_1(T,p)\int_0^T\mathbb{E}\left[\sup_{0\leq\tau\leq s}\|u_\tau^n-u_\tau\|^p\right]ds.
\end{aligned}
$$

By linear growth and Gronwall's inequality, $\mathbb{E}[\sup_{0\leq r\leq t}\|u_r^n\|^p]\leq C_3(T,p)(1+\mathbb{E}[\|\phi(c_0)\|^p]+\int_0^t\mathbb{E}[\|c_s\|^p]ds)$ for some $C_3(T,p)$ that depends on T, p, K_1, and K_2. Then

$$
\begin{aligned}
&\int_0^T 3^{p-1}\mathbb{E}[1+\|u_s^n\|^p+\|c_s\|^p]ds\\
&\leq 3^{p-1}\int_0^T\mathbb{E}[1+\sup_{0\leq r\leq s}\|u_r^n\|^p+\|c_s\|^p]ds\\
&\leq 3^{p-1}\int_0^T\mathbb{E}\left[1+C_3(T,p)\left(1+\mathbb{E}[\|\phi(c_0)\|^p]+\int_0^s\mathbb{E}[\|c_\tau\|^p]d\tau\right)+\|c_s\|^p\right]ds\\
&\leq 3^{p-1}\left(T+C_3(T,p)T(1+\mathbb{E}[\|\phi(c_0)\|^p])\right.\\
&\left.+C_3(T,p)\int_0^T\int_0^s\mathbb{E}[\|c_\tau\|^p]d\tau ds+\int_0^T\mathbb{E}[\|c_s\|^p]ds\right)\\
&\leq 3^{p-1}\left(T+C_3(T,p)T(1+\mathbb{E}[\|\phi(c_0)\|^p])\right.\\
&\left.+C_3(T,p)\int_0^T(T-s)\mathbb{E}[\|c_s\|^p]ds+\int_0^T\mathbb{E}[\|c_s\|^p]ds\right)\\
&\leq 3^{p-1}\left(T+C_3(T,p)T(1+\mathbb{E}[\|\phi(c_0)\|^p])\right.\\
&\left.+C_3(T,p)T\int_0^T\mathbb{E}[\|c_s\|^p]ds+\int_0^T\mathbb{E}[\|c_s\|^p]ds\right)\\
&\leq 3^{p-1}T\left(1+C_3(T,p)(1+\mathbb{E}[\|\phi(c_0)\|^p])\right.
\end{aligned}
$$

$$+C_3(T,p)\sup_{0\le s\le T}\mathbb{E}[\|c_s\|^p]+\sup_{0\le s\le T}\mathbb{E}[\|c_s\|^p]\Bigg)$$
$$=: C_4(T,p).$$

Let $Y(t)=\mathbb{E}[\sup_{0\le r\le t}\|u_r^n-u_r\|^p]$. Then $Y(t)\le\alpha+C_1(T,p)\int_0^t Y(s)ds$, where $\alpha=2^{p-1}T^pL_h^p\mathbb{E}[\sup_{0\le s\le T}\|\theta_s^n-\theta_s\|^p]$. By Gronwall, $Y(T)\le\alpha e^{C_1(T,p)T}$. Since $\theta^n\to\theta$, $\Phi_T(c_0;\theta^n)\to\Phi_T(c_0;\theta)$.

Step 4: Joint Continuity. Let $(c_0^n,\theta^n)\to(c_0,\theta)$. Let $u_t^{n,n}$ be the solution with (c_0^n,θ^n), u_t^n be the solution with (c_0,θ^n), and u_t be the solution with (c_0,θ). Then

$$\mathbb{E}\left[\sup_{0\le t\le T}\|u_t^{n,n}-u_t\|^p\right]\le 2^{p-1}\mathbb{E}\left[\sup_{0\le t\le T}\|u_t^{n,n}-u_t^n\|^p\right]+2^{p-1}\mathbb{E}\left[\sup_{0\le t\le T}\|u_t^n-u_t\|^p\right].$$

By Steps 2 and 3, both terms on the right-hand side go to 0 as $n\to\infty$. Thus, Φ_T is jointly continuous. *Q.E.D.*

Proof of Proposition 4: Let $c^n\to c$ in $L^2(\Omega,\mathcal{F}_t;\mathbb{R}^n)$ and $\theta^n\to\theta$ in $L^2([t,T]\times\Omega;\Theta)$. We want to show that $V_t(c^n;\theta^n)\to V_t(c;\theta)$ in $L^2(\Omega,\mathcal{F}_t;\mathbb{R})$. Let (u_s^n,V_s^n) and (u_s,V_s) be the solutions to the FBSDE system (Forward Propagation)-(Backward Propagation) corresponding to (c^n,θ^n) and (c,θ), respectively, for $s\in[t,T]$. By the stability properties of FBSDEs (see, e.g., Peng and Wu (1999)), we have:

$$\mathbb{E}\left[\sup_{t\le s\le T}\|u_s^n-u_s\|^2+\int_t^T|V_s^n-V_s|^2ds\right]\le C\mathbb{E}\left[\int_t^T\|c_s^n-c_s\|^2ds+\int_t^T\|\theta_s^n-\theta_s\|^2ds\right],$$

for some constant $C>0$ that depends on the Lipschitz constants and time horizon T. Since $c^n\to c$ in L^2 and $\theta^n\to\theta$ in L^2, the right-hand side goes to zero as $n\to\infty$. Thus, $u_s^n\to u_s$ in L^2 uniformly in s, and $V_s^n\to V_s$ in L^2. Now, consider the difference in the value functions:

$$|V_t(c^n;\theta^n)-V_t(c;\theta)|=\left|\mathbb{E}\left[\int_t^T\left(f_{\theta_s^n}(c_s^n,u_s^n,V_s^n)-\lambda\mathcal{L}(u_s^n,\theta_s^n)\right)ds\middle|\mathcal{F}_t\right]\right.$$

$$
\begin{aligned}
&\quad - \mathbb{E}\left[\int_t^T \left(f_{\theta_s}(c_s, u_s, V_s) - \lambda \mathcal{L}(u_s, \theta_s)\right) ds \middle| \mathcal{F}_t\right]\Bigg| \\
&\leq \mathbb{E}\left[\int_t^T \left|f_{\theta_s^n}(c_s^n, u_s^n, V_s^n) - f_{\theta_s}(c_s, u_s, V_s)\right| ds \middle| \mathcal{F}_t\right] \\
&\quad + \lambda \mathbb{E}\left[\int_t^T \left|\mathcal{L}(u_s^n, \theta_s^n) - \mathcal{L}(u_s, \theta_s)\right| ds \middle| \mathcal{F}_t\right].
\end{aligned}
$$

Using the Lipschitz continuity of f_{θ_t} and $\mathcal{L}$, we can bound the terms on the right-hand side by:

$$
\begin{aligned}
&\mathbb{E}\left[\int_t^T \left|f_{\theta_s^n}(c_s^n, u_s^n, V_s^n) - f_{\theta_s}(c_s, u_s, V_s)\right| ds \middle| \mathcal{F}_t\right] \\
&\leq L_f \mathbb{E}\left[\int_t^T \left(\|c_s^n - c_s\| + \|u_s^n - u_s\| + |V_s^n - V_s|\right) ds \middle| \mathcal{F}_t\right], \\
&\lambda \mathbb{E}\left[\int_t^T \left|\mathcal{L}(u_s^n, \theta_s^n) - \mathcal{L}(u_s, \theta_s)\right| ds \middle| \mathcal{F}_t\right] \\
&\leq \lambda L_{\mathcal{L}} \mathbb{E}\left[\int_t^T \left(\|u_s^n - u_s\| + \|\theta_s^n - \theta_s\|\right) ds \middle| \mathcal{F}_t\right].
\end{aligned}
$$

Since $c^n \to c$, $u^n \to u$, $V^n \to V$, and $\theta^n \to \theta$ in L^2, the right-hand sides of these inequalities go to zero as $n \to \infty$. Therefore, $V_t(c^n; \theta^n) \to V_t(c; \theta)$ in $L^2(\Omega, \mathcal{F}_t; \mathbb{R})$. *Q.E.D.*

Proof of Proposition 5: Let $c_t^1 \geq c_t^2$ for all $t \in [0, T]$, and $\phi(c_0^1) \geq \phi(c_0^2)$. Fix a control process θ. Let (u_t^1, V_t^1) and (u_t^2, V_t^2) be the solutions to the FBSDE system (Forward Propagation)-(Backward Propagation) corresponding to c^1 and c^2, respectively. We first show that $u_t^1 \geq u_t^2$ for all t. Since h_{θ_t} is non-decreasing in c, we have $h_{\theta_t}(u, c^1) \geq h_{\theta_t}(u, c^2)$ for any u. By the comparison theorem for SDEs (see, e.g., Ikeda and Watanabe (2014)), and since $\phi(c_0^1) \geq \phi(c_0^2)$, we have $u_t^1 \geq u_t^2$ for all t. Next, we show that $V_t^1 \geq V_t^2$ for all t. Since f_{θ_t} is non-decreasing in c and u, we have $f_{\theta_t}(c_t^1, u_t^1, V) \geq f_{\theta_t}(c_t^2, u_t^2, V)$ for any V. By the comparison theorem for BSDEs

(see, e.g., Pardoux and Peng (1990)), we have $V_t^1 \geq V_t^2$ for all t. Finally, we have:

$$V_t(c^1;\theta) = \mathbb{E}\left[\int_t^T \left(f_{\theta_s}(c_s^1, u_s^1, V_s^1) - \lambda \mathcal{L}(u_s^1, \theta_s)\right) ds \middle| \mathcal{F}_t\right]$$
$$\geq \mathbb{E}\left[\int_t^T \left(f_{\theta_s}(c_s^2, u_s^2, V_s^2) - \lambda \mathcal{L}(u_s^2, \theta_s)\right) ds \middle| \mathcal{F}_t\right] = V_t(c^2;\theta),$$

where the inequality follows from $c^1 \geq c^2$, $u^1 \geq u^2$, $V^1 \geq V^2$, and the monotonicity of f_{θ_t} and $-\mathcal{L}$. Since this holds for any θ, we have $V_t^*(c^1) = \sup_\theta V_t(c^1;\theta) \geq \sup_\theta V_t(c^2;\theta) = V_t^*(c^2)$. Thus, $c^1 \succsim_t c^2$. For strict monotonicity, if there exists an interval $I \subseteq [t,T]$ with positive measure such that $c_s^1 > c_s^2$ for all $s \in I$, and f_{θ_t} is strictly increasing in c over that interval, then the inequality in the above chain becomes strict, and we have $V_t^*(c^1) > V_t^*(c^2)$, which implies $c^1 \succ_t c^2$.

Q.E.D.

Proof of Proposition 6: Fix $\lambda \in [0,1]$, $c^1, c^2 \in \mathcal{D}$, and a control process θ. Let $c^\lambda = \lambda c^1 + (1-\lambda)c^2$. Let (u^i, V^i) solve the FBSDE system (Forward Propagation)-(Backward Propagation) corresponding to c^i for $i \in \{1,2,\lambda\}$. Since ϕ, h_θ, and σ^h are linear in (u,c), $\tilde{u} := \lambda u^1 + (1-\lambda)u^2$ solves the same SDE as u^λ with the same initial condition $\tilde{u}_0 = u_0^\lambda$. Thus, by uniqueness, $u^\lambda = \lambda u^1 + (1-\lambda)u^2$. Similarly, define $\tilde{V} := \lambda V^1 + (1-\lambda)V^2$. Since $\mathcal{L}$ is convex in u, we have $\mathcal{L}(u^\lambda, \theta) \leq \lambda \mathcal{L}(u^1, \theta) + (1-\lambda)\mathcal{L}(u^2, \theta)$. Using this and the concavity of f_θ in (c, V), we get

$$\begin{aligned} -d\tilde{V}_t &= \lambda[f_{\theta_t}(c_t^1, u_t^1, V_t^1) - \lambda \mathcal{L}(u_t^1, \theta_t)]dt - \lambda \sigma_t^1 d\mathcal{Z}_t \\ &\quad + (1-\lambda)[f_{\theta_t}(c_t^2, u_t^2, V_t^2) - \lambda \mathcal{L}(u_t^2, \theta_t)]dt - (1-\lambda)\sigma_t^2 d\mathcal{Z}_t \\ &\geq [f_{\theta_t}(c_t^\lambda, u_t^\lambda, \tilde{V}_t) - \lambda \mathcal{L}(u_t^\lambda, \theta_t)]dt - (\lambda \sigma_t^1 + (1-\lambda)\sigma_t^2) d\mathcal{Z}_t. \end{aligned}$$

Thus, $\tilde{V}$ is a supersolution to the BSDE for V^λ. By the comparison theorem for BSDEs, $V^\lambda \geq \tilde{V} = \lambda V^1 + (1-\lambda)V^2$. Then, using the concavity of f_θ in (c, u, V):

$$V_t(c^\lambda;\theta) = \mathbb{E}_t\left[\int_t^T \left(f_{\theta_s}(c_s^\lambda, u_s^\lambda, V_s^\lambda) - \lambda \mathcal{L}(u_s^\lambda, \theta_s)\right) ds\right]$$

$$\geq \mathbb{E}_t\left[\int_t^T \left(\lambda f_{\theta_s}(c_s^1, u_s^1, V_s^1) + (1-\lambda) f_{\theta_s}(c_s^2, u_s^2, V_s^2) - \lambda \mathcal{L}(u_s^\lambda, \theta_s)\right) ds\right]$$

$$\geq \lambda V_t(c^1;\theta) + (1-\lambda) V_t(c^2;\theta).$$

Since this holds for any θ, we have $V_t^*(c^\lambda) \geq \lambda V_t^*(c^1) + (1-\lambda) V_t^*(c^2)$. Thus, $V_t^*(c)$ is concave in c. *Q.E.D.*

Proof of Proposition 7: Let (c^*, θ^*) be the joint optimal process starting from time t, solving (19). Suppose, for the sake of contradiction, that there exists some $s \in (t, T]$ such that the continuation of (c^*, θ^*) from time s onwards is not optimal for the problem starting at time s. That is, there exists another admissible pair $(\tilde{c}, \tilde{\theta})$ such that:

$$\mathbb{E}_s\left[\int_s^T \left(f_{\tilde{\theta}_\tau}(\tilde{c}_\tau, \tilde{u}_\tau, \tilde{V}_\tau) - \lambda \mathcal{L}(\tilde{u}_\tau, \tilde{\theta}_\tau)\right) d\tau \middle| \mathcal{F}_s\right]$$

$$> \mathbb{E}_s\left[\int_s^T \left(f_{\theta_\tau^*}(c_\tau^*, u_\tau^*, V_\tau^*) - \lambda \mathcal{L}(u_\tau^*, \theta_\tau^*)\right) d\tau \middle| \mathcal{F}_s\right],$$

where $(\tilde{u}, \tilde{V})$ is the solution to the FBSDE system corresponding to $(\tilde{c}, \tilde{\theta})$, with the initial condition $\tilde{u}_s = u_s^*$. Now, consider a new control process $(\hat{c}, \hat{\theta})$ defined as follows:

$$(\hat{c}_\tau, \hat{\theta}_\tau) = \begin{cases} (c_\tau^*, \theta_\tau^*) & \text{if } t \leq \tau < s \\ (\tilde{c}_\tau, \tilde{\theta}_\tau) & \text{if } s \leq \tau \leq T \end{cases}$$

Let $(\hat{u}, \hat{V})$ be the solution to the FBSDE system corresponding to $(\hat{c}, \hat{\theta})$. By the definition of $(\hat{c}, \hat{\theta})$, we have $\hat{u}_s = u_s^*$. By the dynamic programming principle, we have:

$$V_t = \sup_{c,\theta} \mathbb{E}_t\left[\int_t^s \left(f_{\theta_\tau}(c_\tau, u_\tau, V_\tau) - \lambda \mathcal{L}(u_\tau, \theta_\tau)\right) d\tau\right.$$

$$\left. + \mathbb{E}_s\left[\int_s^T \left(f_{\theta_\tau}(c_\tau, u_\tau, V_\tau) - \lambda \mathcal{L}(u_\tau, \theta_\tau)\right) d\tau \middle| \mathcal{F}_s\right] \middle| \mathcal{F}_t\right].$$

Since (c^*, θ^*) is optimal for the problem starting at time t, we have:

$$V_t = \mathbb{E}_t \left[\int_t^s \left(f_{\theta_\tau^*}(c_\tau^*, u_\tau^*, V_\tau^*) - \lambda \mathcal{L}(u_\tau^*, \theta_\tau^*) \right) d\tau \right.$$
$$\left. + \mathbb{E}_s \left[\int_s^T \left(f_{\theta_\tau^*}(c_\tau^*, u_\tau^*, V_\tau^*) - \lambda \mathcal{L}(u_\tau^*, \theta_\tau^*) \right) d\tau \middle| \mathcal{F}_s \right] \middle| \mathcal{F}_t \right].$$

However, for the control process $(\hat{c}, \hat{\theta})$, we have:

$$\begin{aligned}
&\mathbb{E}_t \left[\int_t^s \left(f_{\hat{\theta}_\tau}(\hat{c}_\tau, \hat{u}_\tau, \hat{V}_\tau) - \lambda \mathcal{L}(\hat{u}_\tau, \hat{\theta}_\tau) \right) d\tau \right. \\
&\left. + \mathbb{E}_s \left[\int_s^T \left(f_{\hat{\theta}_\tau}(\hat{c}_\tau, \hat{u}_\tau, \hat{V}_\tau) - \lambda \mathcal{L}(\hat{u}_\tau, \hat{\theta}_\tau) \right) d\tau \middle| \mathcal{F}_s \right] \middle| \mathcal{F}_t \right] \\
&= \mathbb{E}_t \left[\int_t^s \left(f_{\theta_\tau^*}(c_\tau^*, u_\tau^*, V_\tau^*) - \lambda \mathcal{L}(u_\tau^*, \theta_\tau^*) \right) d\tau \middle| \mathcal{F}_t \right] \\
&\quad + \mathbb{E}_t \left[\mathbb{E}_s \left[\int_s^T \left(f_{\tilde{\theta}_\tau}(\tilde{c}_\tau, \tilde{u}_\tau, \tilde{V}_\tau) - \lambda \mathcal{L}(\tilde{u}_\tau, \tilde{\theta}_\tau) \right) d\tau \middle| \mathcal{F}_s \right] \middle| \mathcal{F}_t \right] \\
&> \mathbb{E}_t \left[\int_t^s \left(f_{\theta_\tau^*}(c_\tau^*, u_\tau^*, V_\tau^* - \lambda \mathcal{L}(u_\tau^*, \theta_\tau^*) \right) d\tau \middle| \mathcal{F}_t \right] \\
&\quad + \mathbb{E}_t \left[\mathbb{E}_s \left[\int_s^T \left(f_{\theta_\tau^*}(c_\tau^*, u_\tau^*, V_\tau^*) - \lambda \mathcal{L}(u_\tau^*, \theta_\tau^*) \right) d\tau \middle| \mathcal{F}_s \right] \middle| \mathcal{F}_t \right] = V_t,
\end{aligned}$$

where the strict inequality follows from the assumption that $(\tilde{c}, \tilde{\theta})$ yields a higher value than the continuation of (c^*, θ^*) from time s onwards. This contradicts the optimality of (c^*, θ^*) for the problem starting at time t. Therefore, the continuation of (c^*, θ^*) from time s onwards must be optimal for the problem starting at time s.

Q.E.D.

Proof of Axiom 1: The preference relation $\succsim_t$ is defined by $c \succsim_t c'$ if and only if $V_t^*(c) \geq V_t^*(c')$, where $V_t^*(c)$ is the optimal value function at time t given consumption process c. Since $V_t^*(c)$ is a real-valued function, for any two consumption processes c and c' in $\mathcal{D}$, we must have either $V_t^*(c) \geq V_t^*(c')$ or $V_t^*(c') \geq V_t^*(c)$ (or both, if they are equal). If $V_t^*(c) \geq V_t^*(c')$, then by definition, $c \succsim_t c'$. If $V_t^*(c') \geq V_t^*(c)$, then by definition, $c' \succsim_t c$. Therefore, for any two consumption processes c and c',

we have either $c \succsim_t c'$ or $c' \succsim_t c$ (or both), which means that the preference relation $\succsim_t$ is complete. To show transitivity, suppose $c \succsim_t c'$ and $c' \succsim_t c''$. By definition, this means $V_t^*(c) \geq V_t^*(c')$ and $V_t^*(c') \geq V_t^*(c'')$. Since V_t^* is a real-valued function, the usual transitivity property of the "greater than or equal to" relation for real numbers applies. Therefore, we have $V_t^*(c) \geq V_t^*(c'')$, which implies $c \succsim_t c''$. Thus, the preference relation $\succsim_t$ is transitive.

Q.E.D.

Proof of Axiom 2: Suppose $c_s \geq c'_s$ for all $s \in [t, T]$. By Proposition 5, we have $V_t^*(c) \geq V_t^*(c')$, where $V_t^*(c)$ and $V_t^*(c')$ denote the optimal value functions at time t given consumption processes c and c', respectively. Thus, by the definition of the preference relation, $c \succsim_t c'$. If, in addition, there exists a time interval $I \subseteq [t, T]$ with positive measure such that $c_s > c'_s$ for all $s \in I$, and f_{θ_t} is strictly increasing in c over that interval, then we can refine the argument in the proof of Proposition 5. Specifically, we can show that the solution to the backward SDE (Backward Propagation) corresponding to c is strictly greater than the solution corresponding to c' over the interval I. This implies that $V_t^*(c) > V_t^*(c')$, and therefore $c \succ_t c'$.

Q.E.D.

Proof of Axiom 3: Let $\{c^n\}$ be a sequence of consumption processes such that $c^n \to c$ in $\mathcal{D}$, i.e., $\mathbb{E}[\int_0^T \|c_t^n - c_t\|^2 dt] \to 0$ as $n \to \infty$. We want to show that $\succsim_t$ is continuous, i.e., if $c^n \succsim_t c$ for all n, then $c \succsim_t c$, and if $c \succsim_t c^n$ for all n, then $c \succsim_t c$. By the continuity of the value function (Proposition 4), we have $V_t^*(c^n) \to V_t^*(c)$ as $n \to \infty$. Suppose $c^n \succsim_t c$ for all n. Then $V_t^*(c^n) \geq V_t^*(c)$ for all n. Taking the limit as $n \to \infty$, we have $\lim_{n\to\infty} V_t^*(c^n) = V_t^*(c) \geq V_t^*(c)$, which implies $c \succsim_t c$. Suppose $c \succsim_t c^n$ for all n. Then $V_t^*(c) \geq V_t^*(c^n)$ for all n. Taking the limit as $n \to \infty$, we have $V_t^*(c) \geq \lim_{n\to\infty} V_t^*(c^n) = V_t^*(c)$, which implies $c \succsim_t c$. Thus, the preference relation $\succsim_t$ is continuous with respect to the L^2 norm on $\mathcal{D}$. *Q.E.D.*

3. RESOURCE CONSTRAINTS

A cornerstone of the Neural Differential Utility (NDU) framework is the explicit consideration of the costs associated with the formation and adjustment of preferences. We posit that agents face cognitive and affective resource constraints, thereby

inducing cognitive and affective costs when altering their preference parameters (θ_t). These costs, which we refer to as *preference adjustment costs*, fundamentally shape the dynamics of preference formation, leading to gradual adjustments and potentially explaining phenomena such as endogenous risk and ambiguity aversion, time inconsistency, and loss aversion. This contrasts with traditional models that assume preferences are fixed or evolve exogenously without any associated costs.

3.1. *Cognitive, Affective, and Computational Resources*

We distinguish between three types of resources:

1. **Cognitive Resources:** These pertain to the mental capacities required for deliberation, planning, reasoning, and problem-solving. Cognitive resources are engaged when an agent consciously evaluates and adjusts her preferences or makes decisions based on new information or experiences. This aligns with the concept of "mental effort" or "cognitive control" discussed in the psychology literature (e.g., Botvinick et al. (2001), Kahneman (2011)). Examples include analyzing the potential consequences of different choices, updating beliefs about future outcomes, or devising strategies to achieve long-term goals. Cognitive resources are inherently limited, leading agents to employ heuristics and simplifications in complex decision environments (e.g., Simon (1955, 1972)).
2. **Affective Resources:** These relate to the capacity for emotional regulation and management. Affective resources are deployed when an agent attempts to control or modify her emotional responses to stimuli, or when emotional states influence decision-making. This concept is related to the psychological constructs of "emotional intelligence" (e.g., Salovey and Mayer (1990), Mayer et al. (2008)) and "emotion regulation" (e.g., Gross (1998)). Examples include suppressing immediate impulses, coping with disappointment or regret, cultivating positive emotions associated with certain outcomes, or incorporating anticipated emotions into current choices (e.g., Loewenstein (2000)).
3. **Computational Resources:** These refer to the capacity to process information, perform calculations, and execute algorithms necessary for decision-making.

Computational resources are essential for evaluating complex options, forecasting future outcomes, and optimizing choices in dynamic environments. This concept draws upon the literature on bounded rationality (Simon (1955)) and computational economics (e.g., Sargent (1993), Sims (2003), Tesfatsion (2002)). The limitations in computational resources manifest in an agent's inability to perfectly optimize due to the complexity of the decision problem. This complexity can arise from the sheer number of options, the uncertainty surrounding outcomes, or the intertemporal nature of choices. Agents may therefore resort to approximation algorithms, heuristics, or simplified models of the world (e.g., Gigerenzer and Gaissmaier (2011)).

While conceptually distinct, cognitive, affective, and computational resources are likely to interact in complex ways. For instance, cognitive reappraisal, a strategy for regulating emotions by reinterpreting their meaning, draws upon both cognitive and affective capacities (e.g., Gross (2002)). Similarly, computationally intensive tasks may require significant cognitive effort, leading to fatigue and reliance on heuristics. Our model, however, separates these resources for analytical clarity, allowing us to isolate the effects of each type of resource constraint on preference evolution and decision-making.

3.2. *Formalizing Preference Adjustment Costs*

The limited availability of cognitive and affective resources and the costs associated with their deployment are captured through a preference adjustment cost function. We assume that changing the preference parameters θ_t requires the expenditure of these resources, and that this expenditure is reflected in a cost function $\mathcal{L}(u_t, \theta_t)$. This is analogous to the concept of "adjustment costs" in economics (e.g., Eisner (1963), Lucas Jr (1967)), but specifically tailored to the context of preference formation. We model the preference adjustment cost as a function of the current preference state u_t and the magnitude of the change in preference parameters θ_t. Specifically, we consider a quadratic cost function:

$$\mathcal{L}(u_t, \theta_t) = \frac{1}{2}||\theta_t - \theta_0(u_t)||^2, \tag{26}$$

where $\theta_0(u_t)$ represents a baseline or target preference parameterization given the current preference state u_t. This baseline reflects the agent's "natural" tendency for preference evolution in the absence of active intervention, potentially determined by factors such as habit formation or a default preference structure. The quadratic form, similar to quadratic regularization in machine learning (e.g., Goodfellow (2016)), captures the idea that large or rapid adjustments to preferences are more costly than small, gradual adjustments. This is consistent with evidence from psychology suggesting that cognitive and affective change is often effortful and incremental (e.g., Bechara (2004)).

Interpretation and Properties of the Cost Function. The preference adjustment cost function has several important properties and interpretations:

1. **State Dependence:** The cost depends on the current preference state u_t through the baseline parameterization $\theta_0(u_t)$. This captures the idea that the cost of changing preferences may vary depending on the agent's current psychological state and past experiences.
2. **Increasing Marginal Cost:** The quadratic form implies that the marginal cost of preference adjustment is increasing in the magnitude of the change. This reflects the assumption that it becomes increasingly difficult to deviate further from the baseline preference structure.
3. **Regularization:** The cost function acts as a regularizer in the agent's optimization problem, promoting smooth and gradual preference evolution. This aligns with the principle of Occam's razor, where simpler and more stable preference trajectories are preferred, all else being equal.
4. **Relationship to Information Costs:** While distinct from the rational inattention framework (e.g., Verrecchia (1982), Sims (2003)), the preference adjustment cost can be viewed as analogous to information costs. In rational inattention models, agents face costs for acquiring and processing information. Similarly, in NDU, agents face costs for adjusting their preferences, which can be interpreted as a form of internal information processing.

Remark 4—Regularization Term and Preference Adjustment Costs: The regularization term $\mathcal{L}(u_s, \theta_s)$ in (13) plays a crucial role in the NDU model, representing

the cognitive and affective costs associated with adjusting preferences. Here's a more detailed discussion:

1. **Economic Interpretation**: The regularization term captures the idea that changing preferences is costly. These costs can be cognitive, such as the mental effort required to update beliefs, learn new information, or overcome biases. They can also be affective, such as the emotional discomfort associated with changing habits, deviating from social norms, or experiencing preference reversals.
2. **Specific Examples of Adjustment Costs**:
    - **Cognitive Costs**: These costs can be related to information processing constraints, as in rational inattention models. For example, the cost of acquiring and processing information about new consumption goods or updating beliefs about the future can be represented by a term that penalizes large changes in the parameters θ_s.
    - **Affective Costs**: These costs can be related to psychological factors, such as loss aversion, regret, or the discomfort of changing habits. For example, a term that penalizes deviations from a reference preference state or a default parameter configuration can capture the affective cost of changing preferences.
3. **Mathematical Form of the Regularization Term**: The specific mathematical form of $\mathcal{L}(u_s, \theta_s)$ can be chosen to reflect different economic theories of preference adjustment costs. Some possibilities include:
    - **Quadratic Cost**: $\mathcal{L}(u_s, \theta_s) = \|\theta_s - \theta_0\|^2$, where θ_0 is a baseline or default parameter configuration. This form penalizes deviations from the baseline, with larger deviations incurring higher costs. It can be interpreted as representing a cost of deviating from a default or habitual preference structure.
    - **Entropy-Based Cost**: $\mathcal{L}(u_s, \theta_s) = \sum_i \theta_{s,i} \log \theta_{s,i}$, where $\theta_{s,i}$ are the components of θ_s. This form is inspired by rational inattention models and can be interpreted as a cost of reducing uncertainty about the optimal preference parameters.

- **Sparse Regularization**: $\mathcal{L}(u_s, \theta_s) = \sum_i |\theta_{s,i}|$, also known as L1 regularization. This form encourages sparsity in the parameter vector, meaning that only a few parameters are active at any given time. It can be interpreted as representing a cost of complexity, where simpler preference structures are preferred.

4. **Effect on Agent's Behavior**: The regularization term influences the agent's behavior by creating a trade-off between maximizing instantaneous utility and minimizing preference adjustment costs. A larger value of the parameter λ implies a higher cost of adjusting preferences, leading the agent to be more resistant to changes in the preference state. This can result in smoother preference trajectories and a greater degree of habit persistence.

Remark 5—Connections to Variational Preferences and Rational Inattention: While NDU differs fundamentally from variational preferences and rational inattention, it is possible to establish connections in specific cases.

- **Connection to Variational Preferences**: Suppose the preference adjustment cost $\mathcal{L}(u, \theta)$ in NDU is related to the divergence between probability measures. For example, consider a scenario where the agent has a reference probability measure P and alternative models $Q \in \mathcal{Q}$. If the preference adjustment cost is proportional to the Kullback-Leibler divergence between Q and P, i.e., $\mathcal{L}(u, \theta) \propto D_{KL}(Q||P)$, then the NDU objective function (NDU) would include a term similar to the penalty term in (Robustness). In this case, the regularization parameter λ would play a role analogous to the penalty parameter κ in variational preferences, controlling the trade-off between fitting the data and adhering to the reference model.
- **Connection to Rational Inattention**: Suppose the preference state u_t in NDU represents a compressed or noisy representation of the true underlying state of the world X_t. The agent receives a signal S_t about X_t, and the precision of this signal is determined by the parameters θ_t. If the preference adjustment cost $\mathcal{L}(u, \theta)$ is related to the mutual information between the signal and the state, i.e., $\mathcal{L}(u, \theta) \propto I(S_t; X_t)$, then the NDU objective function (NDU) would include a term similar to the information cost in (Rational Inattention). In this case,

the regularization parameter λ would be analogous to the information cost parameter λ_I, controlling the trade-off between the accuracy of the preference state and the cost of acquiring information.

These connections suggest that NDU can be viewed as a generalization that encompasses aspects of both variational preferences and rational inattention, depending on the specific form of the preference adjustment cost. However, it is important to note that NDU's primary focus is on learning preference dynamics directly from consumption data, rather than imposing a specific structure based on ambiguity aversion or information processing constraints. Further research is needed to explore these connections in more detail and to understand the conditions under which NDU can be formally linked to these alternative models.

3.3. *Resource Constraints, Preference Evolution, and Computational Rationality*

The presence of adjustment and decision-making costs implies that agents cannot instantaneously optimize their utility, nor can they make perfectly rational decisions in the traditional sense. Instead, they face a trade-off between the benefits of adjusting their preferences and making optimal choices and the costs associated with expending cognitive, affective, and computational resources. This trade-off leads to gradual preference shifts over time and the adoption of boundedly rational decision strategies. Agents incrementally adjust their preference parameters in response to new information, experiences, anticipated future outcomes, and the computational complexity of the decision environment. This gradual adjustment can be likened to the process of learning and adaptation in psychology (e.g., Rescorla and Wagner (1972)) and the concept of "computational rationality" in economics (e.g., Sargent (1993), Gabaix (2019), Gabaix and Graeber (2024)). Computational rationality suggests that agents make decisions that are optimal given their resource constraints, even if these decisions deviate from the predictions of traditional models of perfect rationality.

In summary, the resource constraints in NDU play a crucial role in shaping the dynamics of preference evolution and decision-making. They introduce friction into the optimization process, leading to gradual adjustments, bounded rationality, and

the emergence of behavioral phenomena that are difficult to explain with traditional models that assume stable, exogenously determined preferences and unlimited computational capacity. By explicitly modeling the costs associated with cognitive, affective, and computational adjustments and decision-making, NDU provides a more realistic and nuanced account of how preferences are formed, how decisions are made, and how both evolve over time. This approach aligns with the growing body of evidence from behavioral economics, neuroeconomics (e.g., Kahneman (2011), Glimcher (2013)), and computational economics (e.g., Tesfatsion (2002)).

4. DEEP NEURAL NETWORKS FOR UTILITY REPRESENTATION

Traditional approaches to modeling utility in dynamic economic frameworks often rely on parametric functional forms, such as Constant Relative Risk Aversion (CRRA) utility. While analytically tractable, these specifications impose strong restrictions on the nature of preferences and may fail to capture the richness and heterogeneity observed in empirical data, particularly when dealing with high-dimensional state spaces or non-linear interactions between consumption, state variables, and potentially heterogeneous agent characteristics (e.g., Strotz (1955)). To relax these parametric assumptions and allow for more flexible utility representations, we employ deep neural networks (DNNs) and, specifically, Residual Networks (ResNets, see He et al. (2016)), for estimation of the aggregator function and the evolution of the preference state within the Neural Differential Utility (NDU) framework.

4.1. *Deep Neural Networks*

Deep neural networks (DNNs) are a class of flexible function approximators that have achieved significant success in various fields, including computer vision, natural language processing, and increasingly, economics (see LeCun et al. (2015)). A DNN is composed of an input layer, multiple hidden layers, and an output layer. We adapt the standard DNN notation to align with the NDU framework. Let the input layer receive a vector of features, denoted by $\mathbf{x} \in \mathbb{R}^{n_0}$. In the context of NDU, this input could represent current consumption c_t or the current preference state u_t. Each hidden layer l applies an affine transformation followed by a pointwise

non-linear activation function:

$$\mathbf{h}_l = \sigma_l(\mathbf{W}_l \mathbf{h}_{l-1} + \mathbf{b}_l) \tag{27}$$

where $\mathbf{h}_{l-1} \in \mathbb{R}^{n_{l-1}}$ is the output of the previous layer, $\mathbf{W}_l \in \mathbb{R}^{n_l \times n_{l-1}}$ is a weight matrix, $\mathbf{b}_l \in \mathbb{R}^{n_l}$ is a bias vector, and $\sigma_l : \mathbb{R} \to \mathbb{R}$ is a non-linear activation function (e.g., ReLU, sigmoid, tanh). The output layer produces the final prediction. The universal approximation theorem (e.g., Hornik et al. (1989), Cybenko (1989)) guarantees that a DNN with a single hidden layer and a sufficiently large number of neurons can approximate any continuous function on a compact domain to arbitrary accuracy. However, this theorem does not guarantee learnability or good generalization performance in practice, nor does it provide guidance on the required network size. Practical considerations and empirical evidence suggest that deep networks (networks with multiple hidden layers) often exhibit better generalization performance and can learn more complex representations than shallow networks (see Goodfellow (2016)).

4.2. *Residual Networks (ResNets) for Preference State Dynamics*

Despite their representational power, training very deep neural networks can be challenging due to issues such as vanishing and exploding gradients (e.g., Glorot and Bengio (2010)). Residual Networks (ResNets) address this problem by incorporating "skip connections" that allow information to flow more easily across layers. In the NDU framework, we employ ResNets to model the evolution of the preference state u_t. Recall from (Forward Propagation) that the forward propagation of the preference state in NDU is governed by:

$$du_t = h_{\theta_t}(u_t, c_t)dt + \sigma^h_{\theta_t}(u_t, c_t)d\mathcal{Z}^h_t.$$

We define the drift term $h_{\theta_t}(u_t, c_t)$ as a ResNets:

$$h_{\theta_t}(u_t, c_t) = A_t \eta(B_t u_t + \Gamma_t c_t + b_t), \tag{28}$$

where $\theta_t = (A_t, B_t, b_t, \Gamma_t)$ are the time-varying parameters of the ResNets, with $A_t \in \mathbb{R}^{q\times m}$, $B_t \in \mathbb{R}^{m\times q}$, $b_t \in \mathbb{R}^m$, and $\Gamma_t \in \mathbb{R}^{m\times n}$. The function $\eta : \mathbb{R} \to \mathbb{R}$ is a non-linear activation function applied element-wise. This formulation can be interpreted as a continuous-time analog of a discrete-time ResNets layer. A standard residual block computes:

$$\mathbf{h}_{l+1} = \mathbf{h}_l + \mathcal{F}(\mathbf{h}_l; \mathbf{W}_l) \tag{29}$$

where $\mathcal{F}$ is a residual function, typically implemented as a small neural network.

4.3. *Economic Interpretation of the Activation Function and Residual Block*

Activation functions introduce non-linearity into the ResNets architecture, which is crucial for approximating the complex, dynamic relationship between the preference state and consumption. The choice of activation function can significantly impact the model's ability to represent certain types of economic phenomena. Further details on the mathematical properties of activation functions can be found in Goodfellow (2016). The residual block function $h_{\theta_t}(u_t, c_t)$ in (28) updates the agent's preference state (or, equivalently, marginal utility in the context of our model) based on the current preference state u_t and consumption c_t. This update can be decomposed into a series of operations, each with a specific economic interpretation:

1. **Linear Transformation and Interdependencies ($B_t u_t + \Gamma_t c_t$)**:
    - **Influence of Past Preferences ($B_t u_t$)**: The term $B_t u_t$ represents a linear transformation of the current preference state vector u_t, where $B_t \in \mathbb{R}^{q\times q}$. The diagonal elements B_t^{ii} can be interpreted as scaling the influence of the current marginal utility of good i, u_t^i, on its future value. This can capture phenomena like habit formation (positive B_t^{ii}) or satiation (negative B_t^{ii}). The off-diagonal elements B_t^{ij} capture the influence of the marginal utility of good j on the marginal utility of good i, allowing us to model complementarity (positive B_t^{ij}) or substitutability (negative B_t^{ij}) between goods.
    - **Influence of Consumption Path ($\Gamma_t c_t$)**: The term $\Gamma_t c_t$ represents the linear influence of consumption path c_t on the preference state. Γ_t determines the

sensitivity of each component of the preference state to the consumption of different goods. For example, a large positive element in Γ_t suggests that the consumption of a particular good strongly increases the marginal utility of another good.

2. **Exogenous Preference Shocks (b_t)**: The bias vector $b_t \in \mathbb{R}^q$ is added element-wise. This represents exogenous factors that shift preferences independently of past consumption or the current preference state. A positive b_t^i can be interpreted as an external stimulus increasing the desirability of good i, potentially modeling the effects of advertising, social influence, or seasonal changes in taste.
3. **Non-linear Transformation ($\eta(B_t u_t + \Gamma_t c_t + b_t)$)**: The activation function $\eta : \mathbb{R} \to \mathbb{R}$ is applied element-wise to the combined linear transformation and bias. This introduces non-linearity, allowing for more flexible and complex preference patterns that cannot be captured by a purely linear model. Different activation functions imply different types of non-linearities, each with a corresponding economic interpretation:
   - **Rectified Linear Unit (ReLU):** $\text{ReLU}(x) = \max(0, x)$. This function implies that marginal utility only increases if the combined influence of past preferences, current consumption, and exogenous factors is positive. Economically, this could represent a situation where a certain threshold level of a combined "preference signal" is needed before marginal utility increases. It is computationally efficient but can result in inactive neurons if the input is consistently negative (a phenomenon known as the "dying ReLU problem" in the machine learning literature (e.g., Lu et al. (2020))).
   - **Sigmoid:** $\sigma(x) = \frac{1}{1+e^{-x}}$. This function maps its input to the interval $(0, 1)$. In the context of preferences, it could represent a probabilistic interpretation of marginal utility, where the output represents the probability of a good being preferred. However, the sigmoid function can lead to slow learning due to the saturation of the function for large positive or negative inputs, resulting in small gradients (the "vanishing gradient problem").
   - **Hyperbolic Tangent (tanh):** $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. This function maps its input to the interval $(-1, 1)$. It is similar to the sigmoid but is zero-centered,

which can be beneficial for optimization. Economically, a tanh activation could represent situations where marginal utility can be both positive and negative, with a smooth transition between these states.

4. **Scaling and Relative Importance (A_t)**: Finally, the output vector $A_t \in \mathbb{R}^q$ scales the activated marginal utilities element-wise, represented by the Hadamard product ($A_t \odot \eta(B_t u_t + \Gamma_t c_t + b_t)$). This determines the intensity and speed of the preference adjustment for each good. A large magnitude of A_t^i implies that the marginal utility of good i is highly responsive to changes in the transformed input. The relative magnitudes of the elements in A_t effectively determine the relative importance of different goods in the overall utility function at time t.

The continuous-time limit arises from considering increasingly smaller time steps between layers in the discrete-time model. As these steps approach zero, the discrete-time update rule converges to the continuous-time SDE. This continuous-time perspective allows us to analyze the evolution of preferences smoothly and leverage powerful tools from dynamical systems theory.

4.4. *Forward and Backward Propagation in ResNets*

Let's consider a ResNets with multiple residual blocks. We adapt the standard notation to the NDU setting. For a single residual block as defined in (28), the forward pass can be viewed as a discretization of the continuous-time dynamics of the preference state.

1. **Forward Propagation:** The forward pass through a ResNets involves iterating through the residual blocks. For a network with L layers, the forward pass can be written as:

$$\mathbf{h}_1 = \sigma_1(\mathbf{W}_1\mathbf{x} + \mathbf{b}_1)$$
$$\mathbf{h}_{l+1} = \mathbf{h}_l + \mathcal{F}(\mathbf{h}_l; \mathbf{W}_l), \quad l = 1, \ldots, L-1,$$

where $\mathbf{x}$ is the input, which in our case can be c_t or u_t.

2. **Backward Propagation:** For backpropagation, we need to compute the gradients of the loss function with respect to the weights and biases. The addition

of the identity mapping in the gradient calculation helps mitigate the vanishing gradient problem because it ensures that the gradient can flow through the network unimpeded. In the context of NDU, the backward propagation is analogous to the optimization process where the agent adjusts the preference parameters θ_t to maximize expected utility.

3. **Update Rule:** After calculating the gradients, the weights and biases are updated using an optimization algorithm, typically a variant of stochastic gradient descent (SGD). In the NDU framework, this corresponds to the agent adjusting the parameters θ_t over time.

REMARK 6: Connection to Deep Q-Networks and Reinforcement Learning. This paper's use of deep neural networks (DNNs) for utility representation shares conceptual and methodological similarities with Deep Q-Networks (DQN, Mnih et al. (2015)) in reinforcement learning (RL, Watkins and Dayan (1992), Sutton and Barto (2018)), both domains leverage DNNs to approximate complex objective functions via learning optimal policies. In RL, an agent maximizes cumulative rewards through interaction with an environment. DQN approximate the action-value function, or Q-function, $Q(s,a)$, representing the expected cumulative reward for taking action a in state s. The Q-function is parameterized by a deep neural network via parameters θ, learned iteratively via the Bellman equation (also see Proposition 7):

$$Q(s_t, a_t; \theta_t) = r_t + \gamma \max_{a_{t+1}, \theta_{t+1}} Q(s_{t+1}, a_{t+1}; \theta_{t+1}), \tag{30}$$

where r_t is the immediate reward, γ is a discount factor, and s_{t+1} is the next state.

5. MEAN-FIELD CONTROL AND FBSDE SYSTEM

This section analyzes an optimal control problem arising in a dynamic economic model where a representative agent's decisions are influenced by a time-varying preference state, $u_t \in \mathbb{R}^n$. The agent's objective is to choose a control trajectory, represented by parameters $\theta_t \in \Theta \subset \mathbb{R}^q$, to maximize utility. The preference state evolves according to a controlled stochastic differential equation (SDE), influenced by an exogenous process $c_t \in \mathbb{R}^k$ and the control θ_t. The value function V_t repre-

sents the agent's maximized utility from time t onwards. We consider the following forward-backward stochastic differential equation (FBSDE) system, which characterizes the optimal control problem:

$$\begin{cases} du_t &= h(c_t, u_t, \theta_t)dt + \sigma(c_t, u_t, \theta_t)d\mathcal{Z}_t, \quad u_0 = \phi(c_0), \\ dV_t &= -f(c_t, u_t, V_t, \theta_t)dt + Z_t d\mathcal{Z}_t, \quad V_T = 0. \end{cases} \tag{31}$$

Here, $\mathcal{Z}_t$ is a martingale, and $\sigma : \mathbb{R}^k \times \mathbb{R}^n \times \Theta \to \mathbb{R}^{n\times d}$ is the diffusion coefficient. The functions $f : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R} \times \Theta \to \mathbb{R}$ and $h : \mathbb{R}^k \times \mathbb{R}^n \times \Theta \to \mathbb{R}^n$ are assumed to be sufficiently smooth. Specifically, we assume that f and h are continuously differentiable with respect to all their arguments, and that f, h are Lipschitz continuous in u, V, and θ, uniformly in c and t. The initial condition for the preference state is given by $u_0 = \phi(c_0)$, where $\phi : \mathbb{R}^k \to \mathbb{R}^n$ is a given function. The terminal time is $T > 0$. Z_t is an adapted process to be determined, representing the sensitivity of the value function to the martingale. The function h governs the drift of the preference state, while f represents the instantaneous utility function, potentially incorporating a cost of deviation from a baseline control θ_0 (e.g., $f(c_t, u_t, V_t, \theta_t) = \tilde{f}(c_t, u_t, V_t) - \frac{\lambda}{2}\|\theta_t - \theta_0\|^2$). To derive the optimality conditions, we introduce the Hamiltonian:

$$H(c, u, V, \theta, p) = f(c, u, V, \theta) + p^T h(c, u, \theta). \tag{32}$$

Here, $p_t \in \mathbb{R}^n$ is the adjoint process associated with the state equation. Applying the stochastic maximum principle for FBSDEs (see, e.g., Yong and Zhou (2012)), the optimal control θ_t^* satisfies:

$$\frac{\partial H}{\partial \theta}(c_t, u_t^*, V_t^*, \theta_t^*, p_t) = 0, \tag{33}$$

almost surely for all $t \in [0, T]$. The adjoint process p_t can be derived using Itô's lemma and comparing terms with the backward SDE for V_t. Assuming V is a function of (t, u_t), i.e., $V(t, u_t)$, and applying Itô's lemma, comparing this with the

given backward SDE for V_t:

$$dV_t = -f(c_t, u_t, V_t, \theta_t)dt + Z_t d\mathcal{Z}_t, \tag{34}$$

defining $p_t = \frac{\partial V}{\partial u} = \nabla_u V(t, u_t)$, and substituting into the equation for Z_t (using the optimal values u_t^* and θ_t^*), we get $Z_t = \sigma(c_t, u_t^*, \theta_t^*)^T p_t$. The adjoint process p_t is determined by the following backward SDE:

$$dp_t = -\left(\frac{\partial f}{\partial u}(c_t, u_t^*, V_t^*, \theta_t^*) + \left(\frac{\partial h}{\partial u}(c_t, u_t^*, \theta_t^*)\right)^T p_t\right) dt + M_t d\mathcal{Z}_t, \quad p_T = 0. \tag{35}$$

where M_t is an adapted process. The system (31), (33), and (35) forms a coupled FB-SDE system. Under standard assumptions on the Lipschitz continuity and boundedness of the derivatives of f, h, and σ, along with compactness and convexity of Θ, existence and uniqueness can be established using fixed-point arguments or the method of continuation (see, e.g., Ma and Yong (2007), Yong and Zhou (2012)).

Remark 7—Economic Interpretation and Dynamic Programming Principle: The optimal control θ_t^* balances the marginal utility gain from adjusting the preference state (captured by $\frac{\partial f}{\partial \theta}$) with its impact on the preference state's drift (reflected in $p_t^T \frac{\partial h}{\partial \theta}$). The adjoint variable p_t represents the shadow price of the preference state, indicating the marginal value of a small change in u_t. The exogenous process c_t can represent various economic factors, such as aggregate shocks or market prices, influencing the agent's preferences and decisions.

Alternatively, the dynamic programming principle can be used, involving the Hamilton-Jacobi-Bellman (HJB) equation. However, the control's influence on both the instantaneous utility (through f) and the evolution of the value function (via its impact on future u) makes applying the classical HJB equation challenging. The coupled FBSDE system captures this feedback loop. The backward equation for V_t ensures consistency between the value function and the optimal control of the forward equation for u_t. A viscosity solution approach might be necessary to handle potential non-smoothness of the value function (e.g., Yong and Zhou (2012)).

REMARK 8—NDU, Neuron Interactions, and Mean-Field Games: The NDU model, through its optimization of preference parameters θ_s within a ResNets structure, provides a novel framework for analyzing interactions among a large population of neurons, which can be interpreted through the lens of mean-field games (MFGs, (e.g., Achdou et al. (2022), Alvarez et al. (2023))). Here, individual neurons within the ResNets are analogous to "agents" in an MFG. Their collective behavior, influenced by the parameter trajectory θ_s, shapes the evolving preference state u_s. The parameters θ_s can be interpreted as encoding the "strategies" of these interacting neurons, influencing their activation patterns and ultimately determining the aggregate preference output.

The aggregation of individual neuron activations through the ResNets layers implicitly defines a "mean-field" term. This term captures the average influence of the neuron population on each individual neuron's dynamics. A potential approximation for this term could be $\frac{1}{N}\sum_{i=1}^{N} a_i$, where a_i represents the activation of neuron i and N the total number of neurons. Note that neuron activation typically depends on the weighted sum of inputs from the previous layer plus a bias, followed by an activation function. This MFG perspective raises important questions about the NDU model: Does the system exhibit multiple equilibria, each corresponding to a distinct set of emergent preference patterns? How does the stability of these equilibria depend on the network's depth or the number of interacting neurons?

Unlike standard MFGs, NDU focuses on controlling the parameter trajectory θ_s that governs neuron interactions, rather than directly manipulating individual neuron activations. Moreover, the non-linear transformations within the ResNets create more complex interactions than those typically found in MFGs. These distinctions suggest that adapting MFG tools to analyze NDU's emergent dynamics could offer valuable insights into how collective neuron behavior gives rise to complex preference formation. This approach offers a new perspective on the relationship between microscopic neuron interactions and macroscopic preference structures.

Co-editor [Name] Surname; will be inserted later handled this manuscript.