Keeping Unanimity Simple: Intertemporal Collective Choice and Dynamic Consistency*

Xiaoxi Li[†] Xiangyu Qu[‡]

Abstract

A critical feature of many collective choices is the presence of uncertainty at each time period that cannot be resolved currently. A great challenge to collective choices in this context arises from the heterogeneity of individual preferences, under which the unanimity principle often leads to dictatorship. This paper shows that there are very intuitive reasons that unanimity principle should apply only to *simple* alternative comparison. We demonstrate that a non-dictatorial dynamic consistent aggregation rule becomes possible when a simple unanimity principle is introduced.

Keywords: Collective choices, Heterogeneous preferences, Simple Unanimity principle, Dynamic Consistency.

1 Introduction

Many of the economic problems are solved collectively. These issues range from household consumption and saving decisions to central bank monetary policy and commonly designed climate policy among governments[1]. Although collective choices are widely observed, economists do not always agree on what behavioral principles should be imposed

*Financial support from the NSFC project [Grant 72273101] is gratefully acknowledged.

†School of Economics and Management, Wuhan University. Email: xiaoxi-li@whu.edu.cn

‡CNRS, Centre d'economie de la Sorbonne. Email: xiangyuqu@gmail.com

[1] For instance, The Katowice package adopted at the United Nations climate conference (COP24) in December 2018 contains common and detailed rules, procedures and guidelines that operationalize the Paris Agreement for all the member countries.

to create a collective preference. This disagreement stems mainly from the complexity of most collective choice problems, which usually consists of two layers. First, the choice problem has a time component. Decisions not only have an impact on the present, but are also critical for the future. For example, the environmental policy determines future environmental conditions, which in turn have a huge impact on individual future well-beings. Second, the choice problem involves uncertainty. That is, the consequences of a choice decision are usually not resolved. For example, it is extremely uncertain how carbon emissions would change the temperature of the earth, which is one of the key determinants of the environment state.

When confronted with a choice problem that involves both time and uncertainty, one great challenge for collective choice comes from the heterogeneity among individual preferences. Heterogeneous preferences refer to the fact that the consumption values, the patience levels, and the probability evaluations of uncertainty vary across individuals. The conflicts in judgment on different parameters exacerbate the difficulty of constructing a reasonable collective preference. Indeed, heterogeneous preferences are widely present in the daily collective choice problems. Needless to say that individuals differ in consumption values. A most typical example of heterogeneity in discount factors comes from the famous Weitzman investigation. Weitzman (2001) found that economists' opinions on which factor level to use for discounting the future payoffs differed greatly and could not be reconciled. This finding was later reconfirmed by Drupp et al. (2018) and others. An example of individual heterogeneity in the probabilistic judgments of uncertainty comes from the experts' probability estimates on in which degree the carbon emissions would change the Earth's temperature. In a survey conducted by Heal and Millner (2018), the probability estimates of changes in the Earth's temperature varied greatly among experts. This variability in probabilistic judgments is persistent and difficult to reconcile across individuals, as those experts come from different professions who use different models and data for their estimates, making it nearly impossible for an agreement with others by correcting or updating their own to reach a unified value.

Therefore, in this paper, we want to explore the problem of how to aggregate heterogeneous preferences in a framework involving both time and uncertainty. Actually, except for a few studies like Pivato (2022), most literature considered only either the time setting or the uncertainty setting. We in particular are interested in the preference aggregation rule that satisfies dynamic consistency, and discuss what behavioral principles ensure that collective choices obey such aggregation rule. Specifically, we consider a dynamic uncertainty framework similar to Johnsen and Donaldson (1985). We assume that each individual preference as well as the collective preference over intertemporal choices are characterized by an *expected geometric discounting utility* function. In other words, we use three parameters only to fully characterize a time preference (both individual and collective), which are, namely, a felicity function, a discount factor, and a probability measure. Here, the dynamic consistency we consider has two levels of meanings. The first implication is that, separately, both individual and collective preferences satisfy the property of dynamic consistency. In fact, most studies consider only this level of dynamic consistency. The implicit assumption, therefore, is that collective choices have full commitment. That is, the *initial* collective preference is an aggregation of the time-zero individual preferences, and the collective preference after a certain history of time and event will be a conditional preference of the initial collective preference. The deficiency of full commitment is obvious, since the conditional preferences of the initial collective preference are usually not consistent with the aggregated preferences based on the conditional individual preferences. In a democracy, a choice based on a time-zero collective preference might be well superseded by a later choice based on an aggregation of updated individual preferences. To avoid such inconsistency, we assign a second layer of dynamic consistency. That is, the conditional preferences based on the initial collective preference should always be consistent with the aggregated collective preferences based on the individual conditional preferences. By requiring the dynamic consistency with these two implications, we, on the one hand, rule out the unnecessary assumption of full commitment and, on the other hand, make the collective conditional preference always remain aggregation-consistent with the individual conditional preferences. At the same time, it is easy to see that it is the second level of dynamic consistency that makes it necessary to consider and portray the link between the individual conditional preferences and the collective conditional preference. Among others, Dietrich (2021) also discusses dynamic consistent aggregation in a two period situation. The dynamic consistency considered there is about only uncertainty, but not time. However, most decision problems in practice are affected by both time and uncertainty, which motivates our discussions about collective decision making that satisfies the dynamic consistency in the context of Johnsen and Donaldson (1985) framework.

One critical contribution of this paper is to provide an intuitive set of behavioral principles to characterize the dynamically consistent aggregation rule, that is, a utilitarian-maximin-geometric (UMG) rule. Specifically, both in time zero and after any possible history, the collective felicity function (*resp*. probability measure) is a weighted (*resp.* geometric) average of the individual felicity functions (*resp*. probability measures). Moreover, the collective discount factor adopts a maximin perspective on individual discount factor weights, which is the outcome of a 'game' between two conflicting forces, *pessimism* and *optimism*. Actually, the maximin rule is general enough to include many aggregation rules as special cases such as maximum, minimum and linear rules. By and large, UMG rule is general and intuitive.

So, what kind of behavioral principles are we suggesting? Before answering this question, it is important to note that the aggregation rule above is not compatible with the unanimity principle, which is ethically superior to all other alternative principles as claimed by Buchanan and Tullock (1962). The unanimity principle requires that when comparing two alternatives, if all individuals prefer the former, so should the collective. In fact, within the Savage (1954) framework, which involves only uncertainty, Hylland and Zeckhauser (1979) and Mongin (1995) point out that the unanimity principle would lead to the dictatorship if the collective adheres to an expected-utility preference. Within the Koopmans (1960) framework, which involves only time, Zuber (2011) and Jackson and Yariv (2015) obtain a similar impossibility result. Jackson and Yariv (2015) (Page 152) even pessimistically states: "From a policy perspective, non-dictatorial collective choices that are rationalizable by some collective utility function either necessitate commitment devices or involve (choice) reversals over time".

Why, then, does this normatively compelling principle lead to dictatorship? The reason, we believe at least partly, is the improper application of the principle. Indeed, the unanimity principle is usually applied to arbitrary pair of alternatives. However, this unrestricted arbitrariness is problematic if the individual preferences are heterogeneous. Gilboa, Samet and Schmeidler (2004) points out that when there are conflicting value judgments among individuals as well as conflicting probability judgments, it is likely that such conflicts will cancel each other out in choice judgements, resulting in coherent preferences among individuals. Once the collective follows the unanimity principle in this situation, it is difficult to compromise among the conflicting individual parameters and thus has to follow one individual preference exclusively, i.e., the dictatorial rule. This kind of unanimity, formed by multiple conflicting preference parameters, is coined *spurious unanimity* by Mongin (1995).

Therefore, we believe that the scope for applying the unanimity principle should be limited. When individuals compare a pair of alternatives that require the consideration of two or more parameters, the unanimity principle should be adjusted. That is, the collective should not consider individual comparisons of alternatives involving multiple parameters. This is because a spurious unanimity is easily developed in this situation. Only when the alternative pairs are influenced by one parameter only should the collective applies the unanimity principle. This is what we propose: the unanimity principle should be applied only to the *simple* comparison. It is noteworthy that the unanimity principle is also used in a simple way when it comes to Harsanyi (1955), in which individuals have the same probability judgment and differ only in the value judgments.

In fact, the restriction of the unanimity principle has a twofold purpose. On one hand, it allows a single collective parameter to be influenced by the individual parameters of its counterparts only but not other dimensions. This makes meaningful aggregation possible. On the other hand, this restriction makes the individual comparison simpler, being influenced by only a single but not multiple parameters. Simplicity helps individuals to accurately and efficiently reflect their judgments on different parameters. Thus, the collective parameters could be aggregated one by one through this heuristic to form the final collective preference. It is worth noting that this heuristic scheme has its psychological basis. As Daniel Kahneman writes in his mega bestseller *Thinking, Fast and Slow*: "When faced with a difficult question, we tend to answer an easier one, often without noticing the substitution." Here, getting individuals to compare alternatives with a mix of multiple parameters is a complex mission. Instead, the simple heuristic is to have individuals simply compare alternatives that are affected by only a single parameter, thereby aggregating the individual parameters separately. Our idea is embodied in two principles: the *simple unanimity* and the *unanimous separability* principles. Together with the dynamic consistency principle, continuity and independence of irrelevant alternatives, they fully characterize the UMG aggregation rule.

Moreover, the heuristics we discuss here are not limited to collective decision making, which can be applied also to individual decision making in certain situations. A growing body of evidence from neuroscientific investigations suggests that individual brains process and aggregate motivation in parallel. In particular, there is evidence that different parts of the brain respond differently to timed rewards. In addition, multiple probability estimates are also formed in the face of ambiguity[2]. From this perspective, when there are multiple discount factors, multiple felicity functions, and multiple probabilities in

[2] Schmeidler (1989) and Gilboa and Schmeidler (1989) suggest axiomatically that individuals may hold multiple priors when the environment is ambiguous. Hsu et al. (2005) provided evidence in the brain-image in Ellsberg (1961) experiments to confirm this idea.

the individual's own mind, the aggregation of heterogeneous preferences that we study will be relatively helpful in resolving individual decisions if it is still in the individual decision maker's interest to make dynamically consistent decisions.

In this paper, we consider also other aggregation rules for discount factors. As Halevy (2015) pointed out, the expected geometric discounting function is not the only intertemporal function that satisfies the dynamic consistency and aggregation can be discussed in a broader class of intertemporal utilities (Millner and Heal (2018*a*)). We therefore consider such extension and discuss how the heuristic approach contributes to the formation of a collective aggregation rule that satisfies the dynamic consistency when the individual and collective preferences obey the expected additive utility functions.

In sum, this paper is centered on what might be called a *generic argument* why unanimity principle should be limited in a simple manner. The basic logic is intuitive and yet, we believe, powerful and general. In a sense, the relentless force of unanimity principle giving rise to the impossibility result under heterogeneous preferences will act toward dissipating it when the very principle is limited in a simple manner.

The paper is structured as follows: in the next section we will establish the model environment. Section 3 will first define the UMG rule, then discuss the principles used to characterize it, and state one of our main results. In section 4, we will discuss the expected additive utility function and the aggregation based on this assumption. Section 5 discusses the related literature. We conclude in Section 6. All proofs are collected in the Appendix.

2 The Model

2.1 Time and Uncertainty

Time is discrete and varies over the infinite horizon $\mathcal{T} = \{0, 1, \dots\}$. At each period $t > 0$, an event ω_t is drawn from a finite set $\mathcal{S}$, following an initial event ω_0. The state space is $\Omega = \mathcal{S}^{\mathcal{T}}$ with the product algebra. A typical state is a $\mathcal{T}$-indexed sequences $\omega = (\omega_0, \omega_1, \omega_2, \dots)$. The t-period history of events is denoted by $\omega^t = (\omega_0, \omega_1, \dots, \omega_t)$[3] and the set of possible t-histories by $\mathcal{S}^t$. $\mathscr{S} = \bigcup_{t>0} \mathcal{S}^t$ is the set of all possible histories.

Consumption in any history lies in the set X, formally a *convex* and *compact* subset of a vector space. A consumption process has a form $c = (c_t)$, where $c_t : \mathcal{S}^t \to X$ for each $t > 0$. The set of all such consumption processes, denoted by $\mathscr{C}$, is a mixture space under the

[3]An event ω^t can be identified with the subset of state space such that $\omega^t = \{\omega' \in \Omega : \omega'^t = \omega^t\}$.

obvious mixture operation. An important subset of $\mathscr{C}$ is $X^{\mathcal{T}}$, referred to as the subset of *consumption streams*. To elaborate, we identify the consumption process $c = (c_t)$ for which each c_t is a constant with an element ℓ in $X^{\mathcal{T}}$. The consumption levels delivered by any consumption stream ℓ depend only on time t but not on the state ω. Thus, a consumption stream involves time but not uncertainty.

Each individual has a preference ordering on $\mathscr{C}$ at any time-event pair represented by (t,ω). Denote by $\succsim$ the time-zero preference and by $\succsim_{t,\omega}$ the latter preference ordering conditional on the information prevailing at (t,ω). We consider the collection of preference orderings $\{\succsim_{t,\omega}\} := \{\succsim_{t,\omega}: (t,\omega) \in \mathcal{T} \times \Omega\}$.

Environments like this, involving time and uncertainty, are the foundations of most modern economic models, especially in macroeconomics and finance. Our primary interest is to study those theoretical individuals whose preferences are both *additive* and *dynamically consistent* over time and across states.

Let us introduce some further terminologies. Say that a felicity function $u : X \to \mathbb{R}$ is *mixture linear* if $u(\alpha x + (1-\alpha)y) = \alpha u(x) + (1-\alpha)u(y)$ for all x, y in X and $0 \leq \alpha \leq 1$. Say that $u : X \to \mathbb{R}$ is *bounded* if $\max_X u$ and $\min_X u$ both exist in $\mathbb{R}$. Say that a measure p on Ω has *full historical support* if $p(\omega^t) > 0$ for every $\omega^t \in \mathscr{S}$. Let $\mathscr{U}$ and $\mathscr{P}$ denote the set of bounded mixture linear functions and the set of probability measures on Ω with full historical support, respectively.

Though the utility range of c_t for each $t > 0$ is finite for any consumption process c in $\mathscr{C}$, the utility range of c needs not be finite given the infinite horizon. To handle the complication caused by this infinity, we assume that each felicity function is bounded, which guarantees that the representation function defined below is always bounded for all consumption process c. To start with, we assume that the time-zero preference $\succsim$ has the following representation function.

Definition 1. A function W on $\mathscr{C}$ is a *expected geometric discounting utility* (EGDU) if there exists a felicity function $u \in \mathscr{U}$, a discount factor $0 < \delta < 1$ and a probability measure $p \in \mathscr{P}$ such that for all $c \in \mathscr{C}$,

$$W(c) = \mathbb{E}_p\Big[\sum_{t=0}^{\infty} \delta^t u(c_t)\Big]. \tag{1}$$

Let $\mathscr{W}$ denote the set of all EGDU functions. To define the preferences $\succsim_{t,\omega}$, we follow the approach and terminologies of Johnsen and Donaldson (1985) and assume that the preference orderings under time and uncertainty satisfy two different properties. The first condition is the *conditional independence*: the conditional preference $\succsim_{t,\omega}$ depends only on the available information (t,ω), which rules out the possibility that neither any consumption at a time prior to t nor any history event that may have occurred, but did not, has an impact on $\succsim_{t,\omega}$. The second condition is the *dynamic consistency*: for two consumption processes identical up to time t, if the former process is weakly preferred to the latter one in every history at $t+1$, then the former process should be also preferred at (t,ω). Thus, we assume that there exists a felicity function $u \in \mathscr{U}$, a discount factor $0<\delta<1$ and a probability measure $p \in \mathscr{P}$ such that: for every t and ω, $\succsim_{t,\omega}$ is represented by $V_{t,\omega}$, where

$$V_{t,\omega}(c) = \mathbb{E}_{p_{t,\omega}}\Big[\sum_{s\geq t}\delta^{s-t}u(c_s)\Big], \tag{2}$$

in which the conditional probability $p_{t,\omega}$ is defined by, for any $E \subseteq \Omega$,

$$p_{t,\omega}(E) = \frac{p(\{\omega' \in E : \omega'^t = \omega^t\})}{p(\omega^t)} \tag{3}$$

As for the preference orderings defined above, the implied ordering $\succsim_{t,\omega}$ satisfies the conditional independence and the dynamic consistency. Therefore, take a triple $(u,\delta,p) \in \mathscr{U} \times (0,1) \times \mathscr{P}$, the functions $\{V_{t,\omega}\}$ as in (2) are uniquely determined.

In our setting, each $\succsim_{t,\omega}$ can be transformed into a *time-zero* preference under the assumptions of the conditional independence and dynamic consistency. For a t-period history $\omega^t \in \mathcal{S}^t$ and a τ-period history $\omega^\tau \in \mathcal{S}^\tau$, we use (ω^t,ω^τ) to represent the $(t+\tau)$-period history $\bar{\omega}^{t+\tau}$ as a concatenation of histories, i.e. $\bar{\omega}^{t+\tau}_s = \omega_s$ for $s \leq t$ and $\bar{\omega}^{t+\tau}_s = \omega^\tau_{s-t}$ for $s \in [t+1, t+\tau]$. Fix a consumption process $c = (c_\tau)$ and a time-state pair (t,ω), we define a consumption process $c' = (c'_\tau)$ as an *embedding* of c on the event ω^t by: $c'_\tau(\omega^\tau) = c_{t+\tau}(\omega^t,\omega^\tau)$ for any $\tau > 0$ and $\omega^\tau \in \mathcal{S}^\tau$. Fix a probability distribution $p \in \mathscr{P}$ and a time-state pair (t,ω), we define a probability distribution $p' \in \mathscr{P}$ as an *embedding* of p on the event ω^t by: $p'(\omega^\tau) = p_{t,\omega}(\omega^t,\omega^\tau)$ for any $\tau > 0$ and $\omega^\tau \in \mathcal{S}^\tau$.

Definition 2. A collection of functions $\{W_{t,\omega}\}$ is a *generic transformation* of $\{V_{t,\omega}\}$ defined by (u,δ,p) as in (2) if, for each (t,ω), $W_{t,\omega} \in \mathscr{W}$ is determined by $\Big(u,\delta,p'\Big)$ where $p' \in \mathscr{P}$ is an embedding of p on the event ω^t.

In fact, $W_{t,\omega}$ and $V_{t,\omega}$ are generically equivalent in the sense that, for any consumption plan, its estimated value at the present time and its estimated value after the event ω^t is realized are always equal, *i.e.* $W(c') = V_{t,\omega}(c)$ for all $c, c' \in \mathscr{C}$ with c' an embedding of c on the event ω^t. Thereafter, we will work with the *generic* version of preferences defined above.

2.2 Society

A society consists of finite individuals, indexed by $i \in \mathcal{I} = \{1, \dots, n\}$. Each individual i is endowed with a collection of preferences $\{\succsim^i_{t,\omega}\}$ on $\mathscr{C}$, which are generically represented by a collection of time-consistent expected utilities $\{W^i_{t,\omega}\}$.

We assume that a social planner takes a multi-profile approach to determine its preference orderings. That is, a social planner adopts a *preference aggregation rule*, which is a function transforming the individual preference profiles into a social utility form. Formally, a preference aggregation rule is a function $f : \mathscr{W}^n \to \mathscr{W}$. A social planner applies such invariant preference aggregation rule repeatedly at time zero and under each unfolded time-event. At each time-event (t, ω), the social planner collects individual generic conditional preferences $(W^1_{t,\omega}, \dots, W^n_{t,\omega})$ to derive a social generic conditional preference utilising this aggregation rule:

$$W_{t,\omega} = f(W^1_{t,\omega}, \dots, W^n_{t,\omega}).$$

This aggregation procedure reveals two essential conditions we set out. The first condition is *history and future independence*: the social conditional preference at (t, ω) only depends on the individual conditional preferences at (t, ω). That is, the past and future individual preferences do not have an impact on the formation of the current social preference. Since every individual preference is conditionally independent, it is natural that social decisions need not to keep track of the individual behaviors in the past or speculate the individual behaviors in the future. The second condition is *time and event independence*: the social generic function only depends on the individual generic functions, but not on when and under what event these individual functions are obtained. In other words, when faced with the same information, social decisions should be invariant hence not altered by the time when or the manner by which the information is acquired.

3 Dynamic Consistent Aggregation

3.1 UMG Rule

What kind of aggregation rule should the society actually adopt? In this paper, the rule we consider satisfies two important properties that have been recognized in consensus. First, the aggregation rule satisfies the *separability*. It is natural to assume that the instantaneous utility function, discounting function and probability measure of society should depend only on those functions of individuals, respectively. In fact, we do not allow, for example, the social probability to be influenced by the individual patience. Such separability also corresponds to the fact that the preferences of both the society and individuals are composed of these three independent aspects. If we consider that all information and judgments of individuals about events are adequately reflected in the individual probability measures, then the social probability should be limited to individual probability measures only, independent of other factors.

The second property we impose on the aggregation rule is the *consistency*. To make the overall model consistent, the assumptions imposed on individual preferences should be similarly imposed on social preferences. In this sense, the social preference should also be dynamically consistent. This requires that the aggregation function not only translates the individual preference profiles into a social time-additive expected utility function at each time-event pair, but also ensures that the social preference generated by the aggregation function meets the dynamical consistency after the individual preferences are updated in response to new information. In addition, from a normative point of view, it is difficult to imagine a society adopting a dynamically inconsistent updating rule for statistical decision problems. To satisfy these two properties, we propose the following aggregation rule by stating some notations first. Recall that $\mathcal{I} = \{1, \ldots, n\}$. Let $\Delta(\mathcal{I}) = \{\gamma \in [0,1]^n : \sum_{i=1}^n \gamma_i = 1\}$ be the set of all possible weights to each individual. Let $\mathcal{K}(\Delta(\mathcal{I}))$ denote the space of all non-empty closed convex sets of weights, endowed with the Hausdorff topology. A weight-set collection is a non-empty compact collection $\mathbf{\Gamma} \subseteq \mathcal{K}(\Delta(\mathcal{I}))$. Each element $\Upsilon \in \mathbf{\Gamma}$ is a non-empty closed convex set of weights.

Definition 3. A preference aggregation rule $f : \mathscr{W}^n \to \mathscr{W}$ is *Utilitarian-Maximin-Geometric* (UMG) if there exist numbers $\alpha_i \geq 0$ and $\beta_i > 0$ with $\sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i = 1$, and a collection $\mathbf{\Gamma} \in \mathcal{K}(\Delta(\mathcal{I}))$ such that for every $\mathbf{W} \in \mathscr{W}^n$, the social generic expected additive utility $W = f(\mathbf{W})$ satisfies:

- its felicity function is *utilitarian*, i.e. $u = \sum_{i=1}^n \alpha_i u_i$;

- its discount factor is *maximin*, i.e. $\delta = \max_{\Upsilon \in \boldsymbol{\Gamma}} \min_{\gamma \in \Upsilon} \sum_{i=1}^{n} \gamma_i \delta_i$;
- its probability measure is *geometric*, i.e. p on every $\mathcal{S}^t$ is given by $\prod_{i=1}^{n} [p_i(\omega^t)]^{\beta_i}$ up to a multiplication constant.

The UMG rule is indeed a separating rule. In this case, the social felicity function is a weighted sum of individual felicity functions, while the social probability is a weighted product of individual probabilities. The obedience to utilitarianism is a widely used principle for social felicity function, which we will not dwell on. The heuristic for the social probability is to multiply all weighted roots of individual probability values. Obviously, this heuristic satisfies the dynamic consistency, as the society aggregates their posteriors, and the society aggregates their priors first and then applies the updating rule, both yielding the same result. In contrast to other probability aggregating rules, this is compelling because it means that the society behaves like a Bayesian agent - it has a consistent set of priors that are updated according to the Bayes rule. Notice that the maximin aggregation of discount factors is actually quite general, which includes maximum, minimum, and linear rules as special cases. Indeed, if $\boldsymbol{\Gamma} = \left\{ \{\gamma\} : \gamma \in \overline{\Upsilon} \right\}$ for some $\overline{\Upsilon} \subseteq \Delta(\mathcal{I})$, the maximin rule corresponds to $\delta = \max_{\gamma \in \overline{\Upsilon}} \sum_{i=1}^{n} \gamma_i \delta_i$. In particular, if $\overline{\Upsilon}$ includes all the standard basis vectors of $\mathbb{R}^n$, it corresponds to the maximum rule, i.e. $\delta = \max_{i \in \{1,\dots,n\}} \delta_i$. Similarly, if $\boldsymbol{\Gamma} = \{\overline{\Upsilon}\}$, the maximin rule corresponds to $\delta = \min_{\gamma \in \overline{\Upsilon}} \sum_{i=1}^{n} \gamma_i \delta_i$. In particular, if $\overline{\Upsilon}$ includes all the standard basis vectors of $\mathbb{R}^n$, it corresponds to the minimum rule, i.e. $\delta = \min_{i \in \{1,\dots,n\}} \delta_i$. Furthermore, if $\boldsymbol{\Gamma} = \{\overline{\Upsilon}\}$ is a singleton and $\overline{\Upsilon} = \{\gamma\}$ for some $\gamma \in \Delta(\mathcal{I})$ is also a singleton, the maximin rule corresponds to $\delta = \sum_{i=1}^{n} \gamma_i \delta_i$, the linear rule.

3.2 Principles

So what principles do the aggregation functions need to follow in order to be represented as a UMG rule? In this section, we are going to specify these principles that characterize the aggregation rule. In particular, we will carefully discuss the rationality that these principles possess in order to convince the society to accept them.

Here are some notations needed for us to explicitly state the principles. A *felicity profile* is a vector $\mathbf{u} = (u_1, \dots, u_n) \in \mathscr{U}^n$ of individual felicity functions. A *discount profile* is a vector $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n) \in (0,1)^n$ of individual discounting functions. A *probability profile* is a vector $\mathbf{p} = (p_1, \dots, p_n) \in \mathscr{P}^n$ of individual probability measures. A *preference profile* is a vector $\mathbf{W} = (W_1, \dots, W_n) \in \mathscr{W}^n$ of individual preferences.

The first principle formalizes what is usually meant by dynamic consistency.

Dynamic Consistency (DC): For all $\{\mathbf{W}_{t,\omega}\}$ and $\{W_{t,\omega}\}$, if $W_0 = f(\mathbf{W}_0)$, then $W_{t,\omega} = f(\mathbf{W}_{t,\omega})$ for all (t,ω).

This principle assumes that W_0, the aggregation of the time-zero individual preferences $\mathbf{W}_0$, which is followed by those conditional preferences at (t,ω), $W_{t,\omega}$, should yield the same result as the aggregation of individual conditional preferences at (t,ω), $f(\mathbf{W}_{t,\omega})$. In fact, this *consistency* contains two parts. The first part ensures that the aggregated social probability satisfies the property of 'External Bayesianity'. That is, the aggregated social probability at time zero followed by Bayesian updating is always the same as the aggregation of individual updated probabilities. The second part ensures that the aggregation is consistent over time. This means that the aggregated social discount factor remains constant for any given period.

We acknowledge that consistency is not infallible. However, if the principle of consistency is not upheld, social behavior may fall into a certain disorder unless conflicts between social selves at different times and under different events can be convincingly rationalized.

Next we assume that the aggregation function satisfies the property of continuity. This principle is seen as more of a technical requirement, which we do not elaborate further.

Continuity: If a sequence of preference profiles $\mathbf{W}^1,\dots,\mathbf{W}^m,\dots$ pointwisely converges[4] to $\mathbf{W}$ as $m \to \infty$, then $f(\mathbf{W}^1),\dots,f(\mathbf{W}^m),\dots$ pointwisely converges to $f(\mathbf{W})$ as $m \to \infty$.

The next principle restricts the aggregation rule to be separable among felicity, discount factor and probability. It is worth emphasizing that the following axiom of separability was originally proposed by Hylland and Zeckhauser (1979), which is extended into our framework. For any $u \in \mathscr{U}$, let $\underline{u}$ denote vector $\mathbf{u} = (u,\dots,u)$. Similarly, denote $\underline{\delta}$ and $\underline{p}$ the vector $(\delta,\dots,\delta)$ and $(p,\dots,p)$, respectively.

Unanimous Separability (US): There exist functions $g : \mathscr{U}^n \to \mathscr{U}$, $h : (0,1)^n \to (0,1)$ and $k : \mathscr{P}^n \to \mathscr{P}$ such that, for any $\mathbf{W} \in \mathscr{W}^n$ characterized by $(\mathbf{u},\boldsymbol{\delta},\mathbf{p})$, the social lifetime utility $f(\mathbf{W})$ is equipped with $(g(\mathbf{u}),h(\boldsymbol{\delta}),k(\mathbf{p}))$. Furthermore, $g(\underline{u}) = u$, $k(\underline{\delta}) = \delta$ and $k(\underline{p}) = p$ for all $u \in \mathscr{U}, \delta \in (0,1), p \in \mathscr{P}$.

[4]Notice that each $\mathbf{W}^m$ maps a consumption sequence $c \in \mathscr{C}$ to a real vector according to the expected discounted utility summation, Equation (1), the pointwise convergence here refers to the convergence of $\mathbf{W}^m(c)$ to $\mathbf{W}(c)$ as $m \to \infty$ for each $c \in \mathscr{C}$.

Because beliefs about likelihood are irrelevant of evaluations about patience and taste, the social probability should not depend on these two factors. The same reasoning can be applied to the formation of social patience and social tastes as well. This property is imposed by means of different aggregation functions for the three factors. That is, the social felicity function (*resp.* discount factor, or probability measure) is a function only of the individual felicity profile (*resp.* individual discount factor profile, or individual probability profile). Also, US ensures that if the individual felicity utility is the same, the social felicity utility must be consistent with that of the individuals'. The same conclusion applies to the aggregation of discount factors and probability functions as well.

A central principle is *Simple Unanimity*. Some parts of the axiom below are illustrated by parameters to reflect the behavioral meaning in a simple and intuitive way. Indeed, the statement of the behavorial preferences to which they correspond is straightforward.

Simple Unanimity (SU):

(i) if $\mathbf{u}(x) \geq \mathbf{u}(y)$ for $x, y \in X$, then $g(\mathbf{u})(x) \geq g(\mathbf{u})(y)$;

(ii) for all t and all $\omega^t, \omega'^t \in \mathcal{S}^t$, if $\mathbf{p}(\omega^t) \geq \mathbf{p}(\omega'^t)$, then $h(\mathbf{p})(\omega^t) \geq h(\mathbf{p})(\omega'^t)$.

Part (i) in fact responds to the fact that if every individual prefers a certain constant consumption stream, then the society also prefers that consumption stream. Accordingly, it states that if every individual felicity value of an outcome x is higher than that of y, then so is the society.

The second part requires that for any two contemporary historical events, if each individual believes that the probability of the former event is greater than that of the latter, then the society should also follow this probability ranking. It is worth noting that this requirement is different from the standard principle of probabilistic agreement, which requires this principle to hold for any pair of arbitrary events. In contrast, our principle applies only to event pairs of t-period histories (elements in $\mathcal{S}^t$). Moreover, it does not applies to events that, though of the same time, are composed of different historical events before time t. It is easy to see that the principle of probability unanimity without event restrictions is contrary to the geometric probability aggregation rule, while that with event restrictions as in part (ii) is consistent with the geometric probability aggregation rule.

In contrast to the classical one, our SU rules out the possibility that unanimity among individuals is generated by multiple heterogeneity in felicity, discounting and probability functions among individuals. In the phrase of Mongin (1995), our SU avoids the *spurious* *unanimity* that arises from individual heterogeneity.[5]

The last axiom is actually widely applied in the multi-profile aggregation method to establish connections between various preference profiles.

Independence of Irrelevant Alternatives (IIA):

(i) for any $\mathbf{u}, \mathbf{u}'$ and any $x \in X$, if $\mathbf{u}(x) = \mathbf{u}'(x)$, then $g(\mathbf{u})(x) = g(\mathbf{u}')(x)$;

(ii) for any $u, u' \in \mathscr{U}$, $x, y \in X$ and $\boldsymbol{\delta}, \boldsymbol{\delta}'$, if $u(x) + u(y)\delta_i \geq u'(x) + u'(y)\delta_i'$ for all i, then $g(\underline{u})(x) + g(\underline{u})h(\delta) \geq g(\underline{u}')(x) + g(\underline{u}')h(\delta')$;

(iii) for any $\mathbf{p}, \mathbf{p}'$ and any t, if $\dfrac{p_i(\omega^t)}{p_i(\omega'^t)} = \dfrac{p_i'(\omega^t)}{p_i'(\omega'^t)}$ for all i, then $\dfrac{k(\mathbf{p})(\omega^t)}{k(\mathbf{p})(\omega'^t)} = \dfrac{k(\mathbf{p}')(\omega^t)}{k(\mathbf{p}')(\omega'^t)}$.

Part (i) states that if two utility profiles have the same utility evaluation for an alternative, then the social utility, regardless from which utility profile is it aggregated, should have the same utility evaluation for that alternative. This principle reflects the fact that the society is only concerned with the individual utility value of consumption, rather than what the individual utility functions are. In fact, this part has been used by d'Aspremont (1985) and Mongin (1994) to establish the Harsanyi aggregation result in the multi-profile setting.

Part (ii) considers arbitrary pair of preference profiles in which different individuals have the same felicity function. The behavioral implication of part (ii) is that, when considering the consumption behavior in two consequential periods, if the utilities of the first preference profile are greater than that of the second, then the social utility of aggregating the former preference profile will also be greater than aggregating the latter one. In fact, this is the counterpart of part (i) regarding the discount factor. Similarly, this principle also reflects the fact that society is only concerned with the individual consumption utility in the two consequential periods, rather than what the individual discount factors are. It is worth noting the restriction we put on the felicity function. Without it, this axiom would be in a similar predicament as spurious unanimity.

According to the hypothesis of part (iii), the probability ratios under any two history-t events are equal between each pair of individual probabilities in $\mathbf{p}$ and $\mathbf{p}'$. Therefore, the aggregated probabilities for $\mathbf{p}$ and $\mathbf{p}'$ should yield the same ratio under these two events. In fact, if we extend the notion of common-ratio from the pairs of single historical events to that of sets of historical events, it is not difficult to see that the aggregation rule we

[5] We refer to Mongin (1995) for a detailed discussion of spurious unanimity.

consider does not always satisfy such principle. This is the reason why we restrict to single historical events.

To better motivate this principle, consider a situation in which consumption does not change in any period except for period t. Thus, the social planner is only concerned with the impact that the social decision may have on consumption in period t. Specifically, society is faced with two choices. Choice A: consume x if ω^t is realized, and consume z otherwise. Choice B: consume y if the event ω'^t is realized, and consume z otherwise. As mentioned earlier, social choice depends on each individual preference between A and B. Assuming that each individual has exactly the same felicity function u and discount factor δ, the differences among individuals respond only to the likelihood judgment of the events. Thus, each individual preference between A and B depends only on the probability ratio of the two events, ω^t and ω'^t. That is, choice A is preferred to choice B if and only if its probability ratio of ω^t to ω'^t is greater than the ratio of utility differences between y, z and x, z (by assuming felicity of x is strictly higher than that of z). Now consider two different probability profiles $\mathbf{p}, \mathbf{p}'$, if under which each individual has exactly the same probability ratio over all events, then each individual i, regardless of his belief being p_i or p'_i, always has the same choice between A and B. In other words, whether the individual probability profile is $\mathbf{p}$ or $\mathbf{p}'$, the individual preference profile over A and B does not change at all as long as the common-ratio condition holds. Therefore, it is compelling that the social preference between A and B should not change regardless of the individual probability profile being $\mathbf{p}$ or $\mathbf{p}'$. The invariance of social preferences is equivalent to the fact that the social aggregated probabilities also have the common ratio for such event pair.

3.3 Characterization Result

We can now state our main result.

Theorem 1. *A preference aggregation rule $f : \mathscr{W}^n \to \mathscr{W}$ satisfies DC, Continuity, US, SU and IIA if and only if it is UMG.*

The five axioms presented above characterize that the social utility obeys the UMG aggregation rule, thus maintaining dynamic consistency both in time and in uncertainty. It is well known that the geometric mean of individual probabilities is the only probability aggregation rule that makes the social probability remain consistent. However, this uniqueness property does not apply to the maximin aggregation of individual discount factors.

It is not difficult to find that if we simply replace the social discount factor with, say, a geometric aggregation function, (a rule will be discussed extensively in the next section.) the resulting new social function still satisfies the dynamic consistency property. This is, of course, determined by the nature of the social and individual functions as well as the separability of aggregation. Here, we require that the social and individual functions belong to the same domain $\mathscr{W}$. Thus, the individual discount factor profile is always the same regardless of time-event (t, ω). We believe that maximin rule is an appropriate aggregation function since it is general enough to include many popular rules as special cases.

4 Expected Additive Utility Model

In this section, we relax the stationary discounting for EGDU and consider a larger set of preferences. Consider a sequence of time-varying discount factors $\delta(0), \delta(1), \delta(2), \ldots$ where $\delta(t)$ denotes for an agent's discount rate at period t[6]. Say $d : \mathcal{T} \to \mathbb{R}$ is a discount function if d is strictly decreasing and $d(0) = 1$. d is interpreted to be associated with some time-varying discount factor sequence $\delta(0), \delta(1), \delta(2), \ldots$ with $d(t) = \delta(1)\delta(2)\cdots\delta(t)$. Let $\mathscr{D}$ denote the set of discount functions. Now, we assume that the time-zero preference $\succsim$ has the following representation function.

Definition 4. A function $W : \mathscr{C} \to \mathbb{R}$ is an *expected time-additive utility* (ETAU) if there exists a felicity function $u \in \mathscr{U}$, a discount function $d \in \mathscr{D}$, and a probability measure $p \in \mathscr{P}$ such that for all $c \in \mathscr{C}$,

$$W(c) = \mathbb{E}_p\Big[\sum_{t=0}^{\infty} d(t)u(c_t)\Big]. \tag{4}$$

In fact, ETAU includes many popular models as special cases. For instance, if $d(t) = \beta\delta^t$ for $t \geq 1$, then d is a quasi-hyperbolic discounting function. Let $\mathscr{W}^+$ denote the set of expected time-additive utilities. Clearly, $\mathscr{W} \subset \mathscr{W}^+$.

To define the preference $\succsim_{t,\omega}$, we also require that it satisfies two properties, conditional independence and dynamic consistency, as before. However, we do not require the preferences to be *stationary*. Thus, we assume that there exist a felicity function $u \in \mathscr{U}$, a discounting function $d \in \mathscr{D}$ and a probability measure $p \in \mathscr{P}$ such that: for every t and

[6]Discounting a unit of consumption from period t to period $t-1$, for all $t \geq 1$.

ω, $\succsim_{t,\omega}$ is represented by $V_{t,\omega}$, where

$$V_{t,\omega}(c) = \mathbb{E}_{p_{t,\omega}}\Big[\sum_{s\geq t} \frac{d(s)}{d(t)} u(c_s)\Big]. \tag{5}$$

in which the conditional probability $p_{t,\omega}$ is defined as in (3). As for the preference orderings defined above, the implied ordering $\succsim_{t,\omega}$ satisfies conditional independence and dynamic consistency. Therefore, take a triple $(u,d,p) \in \mathscr{U} \times \mathscr{D} \times \mathscr{P}$, the functions $\{V_{t,\omega}\}$ as in (5) are uniquely determined. For the similar reason as we discussed above, each $\succsim_{t,\omega}$ can be transformed into a time-zero preference and, therefore, can be represented by an ETAU.

Definition 5. A collection of functions $\{W_{t,\omega}\}$ in $\mathscr{W}^+$ is a *generic transformation* of $\{V_{t,\omega}\}$ defined by (u,d,p) as in (4) if, for each (t,ω), $W_{t,\omega}$ is determined by $\Big(u, d_t, p'\Big)$, where $d_t(\tau) = \frac{d(\tau+t)}{d(t)}$ for all $\tau \in \mathcal{T}$ and $p' \in \mathscr{P}$ is an embedding of p on the event ω^t.

In the current framework, if the social discount function still adheres to the linear aggregation rule, *i.e.* $k(\mathbf{d}) = \sum_{i=1}^n \beta_i d_i$, then the social function would not satisfy the dynamic consistency property. In fact, at time t, the social discount function $k(\mathbf{d}_t) = \sum_{i=1}^n \beta_i d_{it}$. Time consistency requires that, for $\tau \in \mathcal{T}$,

$$\frac{\sum_{i=1}^n \beta_i d_i(t+\tau)}{\sum_{i=1}^n \beta_i d_i(t)} = \sum_{i=1}^n \beta_i \frac{d_i(t+\tau)}{d_i(t)}$$

The only case for which this equation holds is when certain β_i is one and the rest are zero. To avoid the situation where the social discount function is always determined by a particular individual discount function, while maintaining the dynamic consistency, we consider the following aggregation rule.

Definition 6. A preference aggregation rule $f : \mathscr{W}^{+n} \to \mathscr{W}^+$ is *utilitarian-geometric-geometric* (UGG) if there exist nonnegative numbers α_i, β_i and γ_i with $\sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i = \sum_{i=1}^n \gamma_i = 1$ such that for every $\mathbf{W} \in \mathscr{W}^{+n}$, the social generic utility $W = f(\mathbf{W})$, where

- its felicity function is *utilitarian*, i.e. $u = \sum_{i=1}^n \alpha_i u_i$;
- its discounting function is *geometric*, i.e. $d(t) = \prod_{i=1}^n [d_i(t)]^{\gamma_i}$ for all $t \in \mathcal{T}$;

- its probability measure is *geometric*, i.e. p on every $\mathcal{S}^t$ is given by $\prod_{i=1}^n [p_i(\omega^t)]^{\beta_i}$ up to a multiplication constant.

In order to fully characterize this aggregation rule, we only need to make a few adjustments to the existing axioms. We first need to adjust the US axiom. Obviously, h function is now no longer based on the discount factors, but on the discount functions. We denote $\mathbf{d} \in \mathscr{D}^n$ the discount function profile.

US$^+$: There exist functions $g : \mathscr{U}^n \to \mathscr{U}$, $h : \mathscr{D}^n \to \mathscr{D}$ and $k : \mathscr{P}^n \to \mathscr{P}$ such that, for all $\mathbf{W} \in \mathscr{W}^{+n}$ characterized by $(\mathbf{u}, \mathbf{d}, \mathbf{p})$, the social lifetime utility $f(\mathbf{W})$ is equipped with $(g(\mathbf{u}), h(\mathbf{d}), k(\mathbf{p}))$. Furthermore, $g(u, \ldots, u) = u$, $k(d, \ldots, d) = d$ and $k(p, \ldots, p) = p$ for all $u \in \mathscr{U}, \delta \in (0,1), p \in \mathscr{P}$.

Second, the axiom, IIA(ii), used to describe the maximin aggregation of discount factors will no longer apply since the discount function aggregation rule is geometric. Instead, we impose an alternative axiom.

Say a consumption stream $\ell \in X^{\mathcal{T}}$ is *diperiodic* if there exist $s \in \mathcal{T}$ and $x \in X$ such that $\ell_t = x$ for all $t \in \mathcal{T} \setminus \{0, s\}$. Say two consumption streams $\ell, \ell' \in X^{\mathcal{T}}$ are *co-diperiodic* if there exists $s \in \mathcal{T}$ and $x \in X$ such that $\ell_t = \ell'_t = x$ for all $t \in \mathcal{T} \setminus \{0, s\}$.

SU (iii)$^+$: Let $\mathbf{W} \in \mathscr{W}^+$ be equipped with $\mathbf{u} = (u, \ldots, u)$ and $\mathbf{d} \in \mathscr{D}^n$. For any co-diperiodic consumption streams ℓ, ℓ', if $\mathbf{W}(\ell) \geq \mathbf{W}(\ell')$, then $f(\mathbf{W})(\ell) \geq f(\mathbf{W})(\ell')$.

Theorem 2. *A preference aggregation rule $f : \mathscr{W}^{+n} \to \mathscr{W}^+$ satisfies DC, Continuity, US$^+$, SU(i)(ii)(iii)$^+$ and IIA(i)(iii) if and only if it is UGG.*

In this result, if the individual discount functions are exponential, then the social discount function is also exponential and is a geometric average of individual discount functions, which is different from what we derive in Theorem 1.

5 Related Literature and Discussion

Traditional studies of collective decision making usually assume that the individual preferences satisfy belief homogeneity, e.g., Harsanyi (1955). While this assumption facilitates the aggregation of preferences and hence welfare analysis, it is quite different from the reality. Therefore, most recent studies have turned to the assumption that individual preferences are heterogeneous. As stated in the Introduction, many studies have shown that the heterogeneity assumption takes away the original *magic* of the unanimity principle. In an uncertain setting, Mongin (1995) and Chambers and Hayashi (2006) found the unanimity principle to be incompatible with even Savage's axioms 3 and 4. To break through the impossibility theorem, Gilboa, Samet and Schmeidler (2004) was the first to propose the concept of restricted consistency and proved that the principle is consistent with the linear belief aggregation rule. This aggregation rule was obtained by Billot and Qu (2021*a*) in a more general framework. Dietrich (2021) pointed out that the linear average of probabilities does not necessarily satisfy the dynamic consistency and creatively proposed how to characterize the belief geometric average aggregation rule. However, as pointed out by Pivato (2022), the axiomatization is not done in a dynamic framework, as ours. In this sense, this paper addresses how to characterize the geometric rule in a dynamic framework, and therefore responds to Pivato's concern in a complete way.

It is worth noting that many studies based on the uncertainty framework abandon the assumption of collective expected utility preferences in favor of the collective non-Bayesian preference assumption. This assumption is certainly a reasonable one under ambiguity, or situations where the collective lacks the ability to assign individual weights. Studies based on the non-Bayesian model, like Schmeidler (1989), Gilboa and Schmeidler (1989), include Crès, Gilboa and Vieille (2011), Alon and Gayer (2016), Qu (2017) among many others. However, the belief aggregation rules provided by these studies in general do not satisfy dynamic consistency in the sense of Epstein and Schneider (2003). Therefore, how to characterize a dynamically consistent collective maxmin expected utility will be one of the important topics for future research. Alternatively, Danan et al. (2016) studied the aggregation of incomplete preferences like Bewley (2002). Although their result satisfies the dynamic consistency, incomplete preference assumptions often create difficulties for collective choice. Therefore, how to complete the preference so that it remains dynamically consistent still needs further research.

In the pure time setting, both Zuber (2011) and Jackson and Yariv (2015) found the unanimity principle to be incompatible with time consistency. In the environment where individuals only differ in discount factors, Chambers and Echenique (2018) proposes three rules for aggregating discount factors. One of them suggests aggregation through a weighted average method. This aggregation rule is characterized by Billot and Qu (2021*b*) under heterogeneous preferences. In contrast, Feng and Ke (2018)[7] proposed intergen-

[7]See Bernheim (1989), Farhi and Werning (2007), Caplin and Leahy (2004) for further justification that the collective discount factor should be higher than that of all individuals.

erational unanimity and characterized a constant collective discount factor that is larger than any individual factor. [8] Chichilnisky, Hammond and Stern (2020) argues that future generations are threatened with extinction, and thus proposes a social discounting of extinction. There are many other approaches to study collective time consistency. Similar work, but within a continuous time framework, was studied by Drouhin (2020). However, the biggest difference with this paper is that these above studies do not consider the issue of uncertainty. Therefore, the research methods used are also different from this paper.

Interestingly, Millner and Heal (2018*b*) finds that if we allow collective preferences to not follow the time invariance principle, then we can ensure that collective choices satisfy time consistency by changing the individual weights at different times. This is in contrast to the invariance of weights required in this paper. Similarly, it might be conceivable to make collective choices satisfy dynamic consistency by modifying the weights corresponding to each parameter of each time-event pair in our dynamic framework. However, even if this idea could be implemented, the dynamic consistency it satisfies is not entirely plausible. When the weights are variable, it is easy to imagine that even when two generic profiles are identical, they correspond to completely different aggregation weights. However, the collective needs justifiable reasons beyond the pursuit of dynamic consistency to allow the collective to choose differently in the face of perfectly same generic profiles of individual preferences.

The closest study to this paper is Pivato (2022), in which the framework also incorporates both time and uncertainty. However, Pivato (2022) argues that collectives should only aggregate stable individual parameters. In contrast, when the individual parameter is unstable, the collective does not have to apply any aggregation rule and should instead select that parameter in a discretionary fashion. Since the conditional probabilities of individuals will vary depending on the information realizations, thus Pivato (2022) argues that the probabilities of individuals are unstable. Therefore, the collective should not aggregate the probabilities of individuals. Relatively, since the individual felicity functions are invariant with respect to different information, society should aggregate them. Since the aggregation rule proposed in this paper is separable, our proposed method can still be applied to the case where only certain parameters are aggregated and not others. It is particularly noteworthy that, unlike this paper, the Pivato (2022) model does not include

[8] Drugeon and Wigniolle (2020) studied a similar collective decision problem that was studied by assuming superlinear discounting of individuals.

the discount factor and does not discuss the dynamic consistency of collective choices.

6 Conclusion

The main message conveyed in this paper is that when aggregating heterogeneous preferences of consumption processes in a non-dictatorial way, it is only necessary to respect the unanimity principle in a simple manner (separately), but not in the traditional sense (entirely). The idea of simple unanimity not only effectively avoids spurious unanimity, but also facilitates individuals to accurately convey each parameter value one by one, making non-dictatorial aggregation rule possible. This insight is not only useful for solving aggregation problems, but also for a certain type of individual decision making. Under incomplete preferences, individual preferences can be represented by multiple utility functions (Galaabaatar and Karni (2013)), and our insight may serve individuals to complete their preferences when they have to make certain decisions.

Many of the research (Marglin (1963),Feldstein (1964),Jackson and Yariv (2014)) seem to suggest that policymakers cannot avoid a dynamically inconsistent representative agent. The results of this paper suggest that this frustrating finding is rooted in spurious unanimity principle. Therefore, policymakers should avoid adopting it in the traditional sense and instead respect only simple unanimity principle. In practice, policymakers may not need individuals to compare all possible policies because of the multiple complexities involved. Instead, policymakers might replace them with a series of simple comparisons, as Kahneman suggests. Such comparisons, for instance, may include separately comparing different experts' discount factors, and comparing different experts' probability estimates of events.

APPENDIX

A Preliminary Results

In this appendix we review some established analytic results and notions that are used to prove the results in the main text.

A.1 Functional Analysis

In this subsection we mainly adopt the concepts from Gilboa and Schmeidler (1989). Let L be a non-singleton interval in the real line. $\mathbb{1}$ is the vector in $\mathbb{R}^n$ that has 1 in every coordinate. For any $\phi, \varphi \in \mathbb{R}^n$, we write $\phi \geq \varphi$ (*resp.* $\phi > \varphi$) if $\phi_i \geq \varphi_i$ (*resp.* $\phi_i > \varphi_i$) for each $i = 1, \ldots, n$. For a in $\mathbb{R}$, $a\phi$ is the scalar product, i.e. $(a\phi)_i = a\phi_i$ for each $i = 1, ..., n$, and $\langle \phi, \varphi \rangle = \sum_{i=1}^n \phi_i \varphi_i$ is the inner product of vectors.

A functional $I : L^n \to \mathbb{R}$ is called

- *monotonic* if $I(\phi) \geq I(\varphi)$ for all $\phi, \varphi \in L^n$ with $\phi \geq \varphi$;
- *normalized* if $I(a \cdot \mathbb{1}) = a$;
- *homogeneous of degree 1* (HD1) if $I(a\phi) = aI(\phi)$ for $a > 0$ and $\phi, a\phi \in L^n$;
- *constant additive* if $I(\phi + a \cdot \mathbb{1}) = I(\phi) + a$

A.2 Clark derivatives and differentials

The contents in this subsection are mainly from Ghirardato, Maccheroni and Marinacci (2004). Consider a monotonic, HD1 and constant additive functional $I : L^n \to \mathbb{R}$. Then I is locally Lipschitz. For $\phi \in int(L^n)$ and $\xi \in \mathbb{R}^n$, the *Clarke upper directional derivative* is

$$I^\circ(\phi; \xi) = \limsup_{\varphi \to \phi, t \downarrow 0} \frac{I(\varphi - t\xi) - I(\varphi)}{t}.$$

The *Clark differential* at ϕ is defined by

$$\partial I(\phi) = \{m \in \mathbb{R}^n : \langle m, \xi \rangle \leq I^\circ(\phi; \xi), \quad \forall \xi \in L^n\}.$$

Christensen (1972) shows that if I is locally Lipschitz on L^n, then there exist $K \subseteq int(L^n)$ such that $L^n \setminus K$ has Lebesgue measure zero and I is Gateaux differentiable on K. Let ∇ denote a Gateaux derivative. The next result is Corollary A.5 in Ghirardato, Maccheroni and Marinacci (2004).

Lemma A1. *Suppose a functional I is locally Lipschitz and HD1 on L^n. Then we have*

$$\partial I(0) = \overline{\text{co}} \Big\{ \nabla I(\phi) : \phi \in K \Big\},$$

where $K \subseteq \text{int}(L^n)$ is such that $L^n \setminus K$ has Lebesgue measure zero and I is Gateaux differentiable on K.

We now review two Clarke (1990) results in the manner of Chandrasekher et al. (2022).

Lemma A2 (Theorem 2.5.1 in Clarke (1990))**.** *Suppose $I : L^n \to \mathbb{R}$ is locally Lipschitz. Then there exists $K \subseteq \text{int}(L^n)$ such that $L^n \setminus K$ has Lebesgue measure zero and I is differentiable on K. For each $\varphi \in K$ and $\phi \in \text{int}(L^n)$, we have*

$$\partial I(\phi) = \text{co}\Big\{\lim_n \nabla I(\phi_n) : \phi_n \to \phi, \quad \phi_n \in K\Big\}.$$

Lemma A3 (Theorem 2.8.6 in Clarke (1990))**.** *Suppose functional $I : L^n \to \mathbb{R}$ is given by*

$$I(\cdot) = \sup_{t \in T} I_t(\cdot)$$

for some indexed family of functionals $(I_t)_{t \in T}$ with domain L^n. Assume that there exists some $\alpha > 0$ such that $|I_t(\varphi) - I_t(\xi)| \leq \alpha \|\varphi - \xi\|$ for every $t \in T$ and $\varphi, \xi \in \text{int}(L^n)$. Then for every $\phi \in \text{int}(L^n)$, we have

$$\partial I(\phi) \subseteq \text{co}\Big\{\lim_{i \to \infty} \nabla I_{t_i}(\phi_i) : \phi_i \to \phi, \, t_i \in T, \, I_{t_i}(\phi) \to I(\phi)\Big\}.$$

We now review the Proposition A.3 in Ghirardato, Maccheroni and Marinacci (2004)

Lemma A4. *Assume that $0 \in \text{int}(L^n)$ Let functional $I : L^n \to \mathbb{R}$ be a locally Lipschitz. Then*

1. *If I is HD1, then $\partial I(\phi) \subseteq \partial I(0)$ for all $\phi \in \text{int}(L^n)$.*
2. *If I is monotonic and constant-additive, then $\partial I(\phi) \subseteq \Delta(\mathcal{I})$ for all $\phi \in \text{int}(L^n)$.*

A.3 Boolean Representation

In this subsection, we review the Boolean representation of I, the LEMMA A.6 in Chandrasekher et al. (2022). Let K be the set given by Lemma A1.

Lemma A5. *For each $\phi \in K$, we have*

$$I(\phi) = \max_{\varphi \in K} \inf_{\xi \in \Xi_\varphi} \Big\{I(\xi) + \nabla I(\xi) \cdot (\phi - \xi)\Big\},$$

where $\Xi_\varphi = \Big\{\xi \in K : I(\xi) + \nabla I(\xi) \cdot (\varphi - \xi) \geq I(\varphi)\Big\}$.

B Proof of Theorem 1

We first prove the sufficiency of the theorem. The proof consists of three parts. We first demonstrate that the felicity aggregation function $g : \mathscr{U}^n \to \mathscr{U}$ is *utilitarian*. Then, we show that the discounting aggregation function $h : \mathscr{D}^n \to \mathscr{D}$ is *dual*. Finally, we show that the probability aggregation function $k : \mathscr{P}^n \to \mathscr{P}$ is also *geometric*. Since f is equipped with functions (g, h, k), together we demonstrate that f is a UMG aggregation function.

Part I: We prove the utilitarian aggregation for felicity functions.

Let

$$A = \{\mathbf{a} = \mathbf{u}(x) \mid \mathbf{u} \in \mathscr{U}^n \text{ and } x \in X\}.$$

Lemma B1. *For any* $\mathbf{a}, \mathbf{b} \in A$, *there exist* $\mathbf{u} \in \mathscr{U}^n$ *and* $x, y \in X$ *such that* $\mathbf{u}(x) = \mathbf{a}$ *and* $\mathbf{u}(y) = \mathbf{b}$.

Proof. Let $\mathbf{a} = (a_1, \ldots, a_n), \mathbf{b} = (b_1, \ldots, b_n)$ be in A. Let $\xi \in \mathscr{U}$ be such that, for some $x, y \in X$, $\xi(x) \neq \xi(y)$. The existence of ξ is obvious. Define $\mathbf{u} = (u_1, \ldots, u_n)$ by for each $i \in \mathcal{I}$

$$u_i(z) = a_i \frac{\xi(z) - \xi(y)}{\xi(x) - \xi(y)} + b_i \frac{\xi(z) - \xi(x)}{\xi(x) - \xi(y)} \quad \text{for } z \in X. \tag{6}$$

Clearly, each u_i is mixture linear since ξ is mixture linear. Obviously, $\mathbf{u}$ is bounded. Hence, $\mathbf{u} \in \mathscr{U}^n$. By construction, $\mathbf{u}(x) = \mathbf{a}$ and $\mathbf{u}(y) = \mathbf{b}$. □

Lemma B2. *A is a convex set in $\mathbb{R}^n$.*

Proof. It is clear that $A \subseteq \mathbb{R}^n$. We only demonstrate that A is convex. Let $\mathbf{a}, \mathbf{b} \in A$ and $\alpha \in (0, 1)$. Construct $\mathbf{u} = (u_1, \ldots, u_n) \in \mathscr{U}^n$ as in (6). So, $\mathbf{u}(x) = \mathbf{a}$ and $\mathbf{u}(y) = \mathbf{b}$. The convexity of X implies that $\alpha x + (1-\alpha)y \in X$. Therefore, $\mathbf{u}(\alpha x + (1-\alpha)y) = \alpha\mathbf{u}(x) + (1-\alpha)\mathbf{u}(y) = \alpha\mathbf{a} + (1-\alpha)\mathbf{b} \in A$ by the definition of A. □

Lemma B3. *For any* $\mathbf{u}, \mathbf{u}' \in \mathscr{U}^n$ *and* $x, y \in X$, *if* $\mathbf{u}(x) = \mathbf{u}'(y)$, *then* $g(\mathbf{u})(x) = g(\mathbf{u}')(y)$.

Proof. Let $\mathbf{u}, \mathbf{u}' \in \mathscr{U}^n$ and $x, y \in X$ be such that $\mathbf{u}(x) = \mathbf{u}'(y)$. We know that $\mathscr{U}$ contains all the constant functions. Therefore, there is a $\mathbf{u}'' \in \mathscr{U}^n$ such that $\mathbf{u}''(x) = \mathbf{u}''(y)$. IIA(i) implies that $g(\mathbf{u})(x) = g(\mathbf{u}'')(x)$ and $g(\mathbf{u}')(y) = g(\mathbf{u}'')(y)$. SU(i) implies that $g(\mathbf{u}'')(x) = g(\mathbf{u}'')(y)$. Hence, $g(\mathbf{u})(x) = g(\mathbf{u}')(y)$. □

We define a preference relation $\trianglerighteq$ on A: for $\mathbf{a}, \mathbf{b} \in A$, we say $\mathbf{a} \trianglerighteq \mathbf{b}$ if there exist $\mathbf{u} \in \mathscr{U}^n$ and $x, y \in X$ such that $\mathbf{u}(x) = a, \mathbf{u}(y) = b$ and $g(\mathbf{u})(x) \geq g(\mathbf{u})(y)$. By the previous lemmas, it is straightforward that $\trianglerighteq$ is well-defined. Next, we show that $\trianglerighteq$ satisfies von-Neunman Mongenstein axioms.

Lemma B4. $\trianglerighteq$ *on A satisfies the following properties:*

(i) $\trianglerighteq$ *is a weak order;*

(ii) *For $\mathbf{a}, \mathbf{a}', \mathbf{a}'' \in A$, if $\mathbf{a} \triangleright \mathbf{a}' \triangleright \mathbf{a}''$, then there exist $\lambda, \lambda' \in (0,1)$ such that $\lambda \mathbf{a} + (1-\lambda)\mathbf{a}'' \triangleright \mathbf{a}' \triangleright \lambda' \mathbf{a}' + (1-\lambda')\mathbf{a}''$;*

(iii) *For $\mathbf{a}, \mathbf{a}' \in A$, if $\mathbf{a} \trianglerighteq \mathbf{a}'$, then $\lambda \mathbf{a} + (1-\lambda)\mathbf{a}'' \trianglerighteq \lambda \mathbf{a}' + (1-\lambda)\mathbf{a}''$ for all $\mathbf{a}'' \in A$ and $\lambda \in [0,1]$.*

Proof. **(i)** To see $\trianglerighteq$ is a weak order, we show that it is complete and transitive. Completeness of $\trianglerighteq$ follows from Lemma B1 and the fact that $g(\mathbf{u})$ is complete. To see the transitivity, take $\mathbf{a}, \mathbf{a}', \mathbf{a}'' \in A$ such that $\mathbf{a} \trianglerighteq \mathbf{a}'$ and $\mathbf{a}' \trianglerighteq \mathbf{a}''$. So, there exist $\mathbf{u}, \mathbf{u}' \in \mathscr{U}^n$ and $x, y, x', y' \in X$ such that

$$\mathbf{u}(x) = \mathbf{a} \text{ and } \mathbf{u}(y) = \mathbf{a}' \text{ and } g(\mathbf{u})(x) \geq g(\mathbf{u})(y)$$
$$\mathbf{u}'(x') = a' \text{ and } \mathbf{u}'(y') = a'' \text{ and } g(\mathbf{u}')(x') \geq g(\mathbf{u}')(y').$$

Since $\mathbf{u}(y) = \mathbf{u}'(x') = a'$, Lemma B3 implies $g(\mathbf{u})(y) = g(\mathbf{u}')(x')$. Therefore, $g(\mathbf{u})(x) \geq g(\mathbf{u}')(y')$. By Lemma B1, there exist $\mathbf{u}'' \in \mathscr{U}$ and $x'', y'' \in X$ such that $\mathbf{u}''(x'') = \mathbf{a}$ and $\mathbf{u}''(y'') = \mathbf{a}''$. According to Lemma B3, $g(\mathbf{u}'')(x'') = g(\mathbf{u})(x) \geq g(\mathbf{u}')(y') = g(\mathbf{u}'')(y'')$. Therefore, $\mathbf{a} \trianglerighteq \mathbf{a}''$ by definition.

(ii) Let $\mathbf{a}, \mathbf{a}', \mathbf{a}'' \in A$ be such that $\mathbf{a} \triangleright \mathbf{a}' \triangleright \mathbf{a}''$. So, there exist $\mathbf{u}, \mathbf{u}', \mathbf{u}'' \in \mathscr{U}^n$ and $x, y, x', y', x'', y'' \in X$ such that

$$\mathbf{u}(x) = \mathbf{a} \text{ and } \mathbf{u}(y) = \mathbf{a}'' \text{ and } g(\mathbf{u})(x) > g(\mathbf{u})(y)$$
$$\mathbf{u}'(x') = \mathbf{a} \text{ and } \mathbf{u}'(y') = \mathbf{a}' \text{ and } g(\mathbf{u}')(x') > g(\mathbf{u}')(y')$$
$$\mathbf{u}''(x'') = \mathbf{a}' \text{ and } \mathbf{u}''(y'') = \mathbf{a}'' \text{ and } g(\mathbf{u}')(x'') > g(\mathbf{u}')(y'')$$

By Lemma B3, we have

$$g(\mathbf{u})(x) = g(\mathbf{u}')(x') \text{ and } g(\mathbf{u}')(y') = g(\mathbf{u}')(x'') \text{ and } g(\mathbf{u})(y) = g(\mathbf{u}')(y'').$$

So, $g(\mathbf{u})(x) > g(\mathbf{u}')(y') > g(\mathbf{u})(y)$. Since $g(\mathbf{u})$ is linear, there exists a unique $\lambda^* \in (0,1)$ such that

$$g(\mathbf{u}')(y') = g(\mathbf{u})(\lambda^* x + (1-\lambda^*)y).$$

Therefore, for $\lambda \in (\lambda^*, 1)$ and $\lambda' \in (0, \lambda^*)$, we have

$$g(\mathbf{u})(\lambda x + (1-\lambda)y) > g(\mathbf{u}')(y') > g(\mathbf{u})(\lambda' x + (1-\lambda')y).$$

Linearity of $\mathbf{u}$ implies that $\mathbf{u}(\lambda x + (1-\lambda)y) = \lambda a + (1-\lambda)a''$ and $\mathbf{u}(\lambda' x + (1-\lambda')y) = \lambda' a + (1-\lambda')a''$. Hence, $\lambda a + (1-\lambda)a'' \rhd a' \rhd \lambda' a + (1-\lambda')a''$.

(iii) Let $\mathbf{a}, \mathbf{a}', \mathbf{a}'' \in A$ and $\lambda \in [0,1]$. Suppose that $\mathbf{a} \trianglerighteq \mathbf{a}'$. So, there exist $\mathbf{u} \in \mathscr{U}^n$ and $x, y \in X$ such that $\mathbf{u}(x) = \mathbf{a}, \mathbf{u}(y) = \mathbf{a}'$ and $g(\mathbf{u})(x) \geq g(\mathbf{u})(y)$. Also, there exist $\mathbf{u}', \mathbf{u}'' \in \mathscr{U}^n$ and $x', x'', y', y'' \in X$ such that

$$\begin{aligned} \mathbf{u}'(x') = \mathbf{a} \quad &\text{and} \quad \mathbf{u}'(y') = \mathbf{a}'' \\ \mathbf{u}''(x'') = \mathbf{a}' \quad &\text{and} \quad \mathbf{u}''(y'') = \mathbf{a}''. \end{aligned}$$

So, $\mathbf{u}'(\lambda x' + (1-\lambda)y') = \lambda \mathbf{a} + (1-\lambda)\mathbf{a}''$ and $\mathbf{u}''(\lambda x' + (1-\lambda)y') = \lambda \mathbf{a}' + (1-\lambda)\mathbf{a}''$.

Therefore,

$$\begin{aligned} g(\mathbf{u}')(\lambda x' + (1-\lambda)y') &= \lambda g(\mathbf{u}')(x') + (1-\lambda) g(\mathbf{u}')(y') \\ g(\mathbf{u}'')(\lambda x'' + (1-\lambda)y'') &= \lambda g(\mathbf{u}'')(x'') + (1-\lambda) g(\mathbf{u}'')(y'') \end{aligned}$$

Since $\mathbf{u}(x) = \mathbf{u}'(x')$, $\mathbf{u}(y) = \mathbf{u}''(x'')$ and $\mathbf{u}'(y') = \mathbf{u}''(y'')$, Lemma B3 implies that $g(\mathbf{u}')(x') = g(\mathbf{u})(x)$, $g(\mathbf{u}'')(x'') = g(\mathbf{u})(y)$ and $g(\mathbf{u}')(y') = g(\mathbf{u}'')(y'')$. Hence,

$$g(\mathbf{u}')(\lambda x' + (1-\lambda)y') \geq g(\mathbf{u}'')(\lambda x'' + (1-\lambda)y'').$$

Therefore, $\lambda \mathbf{a} + (1-\lambda)\mathbf{a}'' \trianglerighteq \lambda \mathbf{a}' + (1-\lambda)\mathbf{a}''$. □

Lemma B5. *There exist nonnegative numbers $\alpha_1, \ldots, \alpha_n$ such that, for any $\mathbf{u} \in \mathscr{U}^n$, $g(\mathbf{u})(x) = \sum_i \alpha_i u_i(x)$ for all $x \in X$.*

Proof. We know from the above Lemmas that $\trianglerighteq$ on A satisfies the three properties and A is a convex set. According to, for instance, the mixture space theorem of Herstein and Milnor (1953), there are real numbers $\alpha_0, \alpha_1, \ldots, \alpha_n$ such that $\trianglerighteq$ is represented by a linear function: for all $\mathbf{a} \in A$,

$$\mathbf{a} \longmapsto \sum_i \alpha_i a_i + \alpha_0.$$

Furthermore, by SU(i), it is immediate to see that each α_i is nonnegative for $i \neq 0$. According to the definition of $\trianglerighteq$, we must have for all $\mathbf{u}$ and $x \in X$, if $\mathbf{u}(x) = a$, then $g(\mathbf{u})(x) =$

$\sum_i \alpha_i a_i + \alpha_0 = \sum_i \alpha_i u_i(x) + \alpha_0$. By the US axiom, $g(u,\dots,u) = u$, which implies $\alpha_0 = 0$. $\square$

Part II: We prove the maximin aggregation for discount factors.

Lemma B6. *There exists a monotonic, HD1, constant additive functional $I : \mathbb{R}^n \to \mathbb{R}$ such that for all $\boldsymbol{\delta}, \boldsymbol{\delta}' \in (0,1)^n$*

$$h(\boldsymbol{\delta}) \geq h(\boldsymbol{\delta}') \Longleftrightarrow I(\boldsymbol{\delta}) \geq I(\boldsymbol{\delta}').$$

Proof. We first demonstrate that h is monotonic, HD1 and constant additive on $(0,1)^n$. To see monotonicity, we need to show that $\boldsymbol{\delta} \geq \boldsymbol{\delta}'$ implies $h(\boldsymbol{\delta}) \geq h(\boldsymbol{\delta}')$. Let $u \in \mathscr{U}$ and $x, y \in X$ be such that $u(x) = 0$ and $u(y) = 1$.Therefore, $\delta \geq \delta'$ implies that for each i,

$$u(x) + u(y)\delta_i \geq u(x) + u(y)\delta_i'$$

According to IIA(ii), we have

$$g(\underline{u})(x) + g(\underline{u})(y)h(\boldsymbol{\delta}) \geq g(\underline{u})(x) + g(\underline{u})(y)h(\boldsymbol{\delta}').$$

As demonstrated, we know $g(\underline{u}) = u$. So, it is immediate from the above that $h(\boldsymbol{\delta}) \geq h(\boldsymbol{\delta}')$.

To see h is HD1, let $\boldsymbol{\delta}, \boldsymbol{\delta}' \in (0,1)^n$ and $a \in (0,1)$ be such that $\boldsymbol{\delta} = a\boldsymbol{\delta}'$. Take $u, u' \in \mathscr{U}$ and $x, y \in X$ such that $u(x) = u'(x)$ and $a \cdot u(y) = u'(y) \neq 0$. Therefore, for each i

$$u(x) + u(y)\delta_i = u'(x) + u'(y)\delta_i'$$

By IIA(ii),

$$g(\underline{u})(x) + g(\underline{u})(y)h(\boldsymbol{\delta}) = g(\underline{u}')(x) + g(\underline{u}')(y)h(\boldsymbol{\delta}').$$

Since $g(\underline{u}) = u$ and $g(\underline{u}') = u'$, we have $u(y)h(\boldsymbol{\delta}) = u'(y)h(\boldsymbol{\delta}')$, which is $h(a\boldsymbol{\delta}') = ah(\boldsymbol{\delta}')$.

To see h is constant additive, assume that $a \in \mathbb{R}$ and $\boldsymbol{\delta}, \boldsymbol{\delta} + a\mathbb{1} \in (0,1)^n$. Take $u, u' \in \mathscr{U}$ and $x, y \in X$ such that $u(x) = u'(x) + a$ and $u(y) = u'(y) = 1$. Obviously, for all i,

$$u(x) + u(y)\delta_i = u'(x) + u'(y)(\delta_i + a)$$

Therefore, IIA(ii) implies

$$g(\underline{u})(x) + g(\underline{u})(y)h(\boldsymbol{\delta}) = g(\underline{u}')(x) + g(\underline{u}')(y)h(\boldsymbol{\delta} + a\mathbb{1}).$$

Again, $g(\underline{u}) = u$ and $g(\underline{u}') = u'$ imply that $u(x) + u(y)h(\boldsymbol{\delta}) = u'(x) + u'(y)h(\boldsymbol{\delta} + a\mathbb{1})$, which is $a + h(\boldsymbol{\delta}) = h(\boldsymbol{\delta} + a\mathbb{1})$.

Now, we define functional I. First, if $\mathbf{a} \in (0,1)^n$, define $I(\mathbf{a}) = h(\mathbf{a})$. If $\mathbf{a} \in \mathbb{R}^n_+ \setminus (0,1)^n$, let $a^* = \max_i a_i + 1$. Clearly, $\frac{\mathbf{a}}{a^*} \in (0,1)^n$. So, define $I(\mathbf{a}) = a^* \cdot h(\frac{\mathbf{a}}{a^*})$. For the rest possible $\mathbf{a}$, let $a_* = \min a_i - 1$. Obviously, $\mathbf{a} - a_* > 0$. We, therefore, define $I(\mathbf{a}) = I(\mathbf{a} - a_*) + a_*$. By this definition, it is routine to check that I is monotonic, constant additive and HD1 in $\mathbb{R}^n$. □

The rest of proof is similar to the proof of Theorem 1 in Chandrasekher et al. (2022). Let

$$\Gamma^* := \text{cl}\{\Upsilon^*_\phi : \phi \in \mathbb{R}^n\}, \quad \text{where } \Upsilon^*_\phi = \{\gamma \in \partial I(0) : \langle \gamma, \phi \rangle \geq I(\phi)\}.$$

Write $\gamma_\phi := \nabla I(\phi)$. Since I is locally Lipschitz and HD1, according to Lemma A1, $\nabla I(\phi) \in \partial I(0)$, which is $\gamma_\phi \in \partial I(0)$. Let the set K be such that $\mathbb{R}^n \setminus K$ has Lebesgue measure zero.

Take $\phi \in K$ and $a \in (0,1)$. HD1 of I implies that $I(a\phi) = aI(\phi)$ and $\nabla I(a\phi) = \nabla I(\phi)$. Thus, $I(a\phi)$ is differentiable at each $a \in (0,1)$. Hence,

$$I(\phi) = \left\langle \gamma_\phi, \phi \right\rangle. \tag{7}$$

According to dual representation result, Lemma A5, we have

$$I(\phi) = \max_{\varphi \in K} \inf_{\xi \in \Xi_\varphi} \left\{ I(\xi) + \nabla I(\xi) \cdot (\phi - \xi) \right\},$$

where $\Xi_\varphi = \left\{ \xi \in K : I(\xi) + \nabla I(\xi) \cdot (\varphi - \xi) \geq I(\varphi) \right\}$.

By (7), it can be simplified as following:

$$I(\phi) = \max_{\varphi \in K} \inf_{\xi \in \Xi_\varphi} \left\langle \gamma_\xi, \phi \right\rangle, \tag{8}$$

where $\Xi_\varphi = \left\{ \xi \in K : \left\langle \gamma_\xi, \varphi \right\rangle \geq I(\varphi) \right\}$. Let $\Upsilon_\varphi := \left\{ \gamma_\xi : \xi \in K; \left\langle \gamma_\xi, \varphi \right\rangle \geq I(\varphi) \right\}$. Therefore, we have $\xi \in \Xi_\varphi \Longleftrightarrow \gamma_\xi \in \Upsilon_\varphi$. Also, by Lemma A1, we know $\overline{\text{co}}(\Upsilon_\varphi) = \Upsilon^*_\varphi$. Hence, (8) can be written as

$$I(\phi) = \max_{\varphi \in K} \min_{\gamma \in \Upsilon^*_\varphi} \langle \gamma, \phi \rangle. \tag{9}$$

Pick any ϕ, φ in $\mathbb{R}^n$. Take sequences (ϕ_m) and (φ_m) in K such that $\phi_m \to \phi$ and $\varphi_m \to \varphi$.

For each m, let

$$\gamma_m \in \arg\min_{\gamma \in \Upsilon^*_{\varphi_m}} \langle \gamma, \phi_m \rangle.$$

By the definition of $\Upsilon^*_{\varphi_m}$, we have

$$\langle \gamma_m, \varphi_m \rangle \geq I(\varphi_m).$$

Let γ^* be such that $\gamma_m \to \gamma$. The continuity of I implies that $\langle \gamma^*, \varphi \rangle \geq I(\varphi)$. Hence, we have $\gamma^* \in \Upsilon^*_\varphi$.

Also, by (9), we know

$$\langle \gamma_m, \phi_m \rangle = \min_{\gamma \in \Upsilon^*_{\varphi_m}} \langle \gamma, \phi_m \rangle \leq \left\langle \gamma_{\phi_m}, \phi_m \right\rangle = I(\phi_m).$$

Again, the continuity of I implies that

$$\min_{\gamma \in \Upsilon^*_\varphi} \langle \gamma, \phi \rangle \leq \langle \gamma^*, \phi \rangle \leq I(\phi). \tag{10}$$

Since this inequality holds for arbitrary φ, it is true that for all $\Upsilon \in \mathbf{\Gamma}^*$,

$$\min_{\gamma \in \Upsilon} \langle \gamma, \phi \rangle \leq I(\phi).$$

According to the definition of Υ^*_ϕ, we have

$$\min_{\gamma \in \Upsilon^*_\phi} \langle \gamma, \phi \rangle \geq I(\phi).$$

Hence,

$$\min_{\gamma \in \Upsilon^*_\phi} \langle \gamma, \phi \rangle \leq I(\phi) \leq \min_{\gamma \in \Upsilon^*_\phi} \langle \gamma, \phi \rangle,$$

which is

$$I(\phi) = \min_{\gamma \in \Upsilon^*_\phi} \langle \gamma, \phi \rangle = \max_{\Upsilon \in \mathbf{\Gamma}^*} \min_{\gamma \in \Upsilon} \langle \gamma, \phi \rangle.$$

Part III: We prove the geometric aggregation for probability measures.

Take $E, F \in \mathcal{S}^t$, define function $\varphi^t_{EF} : (0,\infty)^n \to (0,\infty)$ by for $\mathbf{a} \in (0,\infty)^n$, where there is

$\mathbf{p} \in \mathscr{P}^n$ such that $\mathbf{a} = \left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right)$,

$$\varphi^t_{EF}(\mathbf{a}) = \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(F)}.$$

We claim that this function is well-defined. For every t, define

$$\mathscr{B}^t = \left\{\left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right) : \mathbf{p} \in \mathscr{P}^n \text{ and } E, F \in \mathcal{S}^t\right\}.$$

Note that $(0, \infty)^n = \mathscr{B}^t$. Since every $p \in \mathscr{P}$ has locally full support, the above equality is clearly true. To see φ^t_{EF} is well-defined, it suffices to show that for $\mathbf{p}, \mathbf{p}' \in \mathscr{P}^n$,

$$\left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right) = \left(\frac{p'_1(E)}{p'_1(F)}, \ldots, \frac{p'_n(E)}{p'_n(F)}\right) \quad \text{implies} \quad \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(F)} = \frac{k(\mathbf{p}')(E)}{k(\mathbf{p}')(F)}.$$

Notice that $\frac{p_i(E)}{p_i(F)} = \frac{p'_i(E)}{p'_i(F)}$ demonstrates that p_i, p'_i have common-ratio between E and F. Therefore, IIA(iii) requires that $k(\mathbf{p})$ and $k(\mathbf{p}')$ also have common-ratio between E and F, which is $\frac{k(\mathbf{p})(E)}{k(\mathbf{p})(F)} = \frac{k(\mathbf{p}')(E)}{k(\mathbf{p}')(F)}$. Hence, for any $E, F \in \mathcal{S}^t$, function φ^t_{EF} is well-defined as above.

Lemma B7. *For any $E, F, G \in \mathcal{S}^t$ and $\mathbf{p} \in \mathscr{P}^n$,*

$$\varphi^t_{EF}\left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right) \cdot \varphi^t_{FG}\left(\frac{p_1(F)}{p_1(G)}, \ldots, \frac{p_n(F)}{p_n(G)}\right) = \varphi^t_{EG}\left(\frac{p_1(E)}{p_1(G)}, \ldots, \frac{p_n(E)}{p_n(G)}\right).$$

Proof. It is straightforward by the definition.

$$\begin{aligned}
\varphi^t_{EF}\left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right) \cdot \varphi^t_{FG}\left(\frac{p_1(F)}{p_1(G)}, \ldots, \frac{p_n(F)}{p_n(G)}\right) &= \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(F)} \cdot \frac{k(\mathbf{p})(F)}{k(\mathbf{p})(G)} \\
&= \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(G)} \\
&= \varphi^t_{EG}\left(\frac{p_1(E)}{p_1(G)}, \ldots, \frac{p_n(E)}{p_n(G)}\right)
\end{aligned}$$

□

Lemma B8. *For any $E, F, E', F' \in \mathcal{S}^t$ and $\mathbf{p} \in \mathscr{P}^n$, if $\left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right) = \left(\frac{p_1(E')}{p_1(F')}, \ldots, \frac{p_n(E')}{p_n(F')}\right)$, then $\varphi^t_{EF}\left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right) = \varphi^t_{E'F'}\left(\frac{p_1(E')}{p_1(F')}, \ldots, \frac{p_n(E')}{p_n(F')}\right)$.*

Proof. Let $\mathbf{p} \in \mathscr{P}^n$ be such that $\left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right) = \left(\frac{p_1(E')}{p_1(F')}, \ldots, \frac{p_n(E')}{p_n(F')}\right) = \mathbf{a}$. If $E = E'$, then

$p_i(F) = p_i(F')$ for all i. By SU (ii), we have $k(\mathbf{p})(F) = k(\mathbf{p})(F')$. According to the definition, it is immediate that

$$\varphi^t_{EF}(\mathbf{a}) = \varphi^t_{E'F}(\mathbf{a}).$$

Similarly, we have $\varphi^t_{EF}(\mathbf{a}) = \varphi^t_{EF'}(\mathbf{a})$. Therefore,

$$\varphi^t_{EF}(\mathbf{a}) = \varphi^t_{E'F}(\mathbf{a}) = \varphi^t_{E'F'}(\mathbf{a})$$

□

Based on the above result, the value of function φ^t_{EF} is independent of events E, F. Therefore, we can subtract the subscript EF and use function φ^t to represent φ^t_{EF}.

Lemma B9. *For any* $\mathbf{a} \in (0,\infty)^n$, $\varphi^t(\mathbf{a}) = \varphi^s(\mathbf{a})$.

Proof. We only need to consider the case where $t \neq s$. Without loss of generality, assume that $t < s$. Pick $\mathbf{a} \in (0,\infty)^n$. Let $E, F \in \mathcal{S}^t$. Let $E', F' \in \mathcal{S}^s$ be such that there is $\omega \in \Omega$ and $E', F' \in \mathcal{F}_s(\omega)$. Let $\mathbf{p} \in \mathscr{P}^n$ be such that $\left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right) = \left(\frac{p_1(E')}{p_1(F')}, \ldots, \frac{p_n(E')}{p_n(F')}\right) = \mathbf{a}$.

Fix a state $\omega^* = (\omega^*_0, \omega^*_1, \omega^*_2, \ldots)$. Let $E = (\omega^*_0, \ldots, \omega^*_t)$ and $F = (\omega^*_0, \ldots, \omega^*_{t-1}, \omega_t)$, where $\omega^*_t \neq \omega_t$. Let $E' = (\omega^*_0, \ldots, \omega^*_s)$ and $F' = (\omega^*_0, \ldots, \omega^*_{s-1}, \omega_s)$, where $\omega^*_s \neq \omega_s$.

According to updating rule, we have for each i,

$$p_i(E' \mid s-t, \omega^*) = \frac{p(\{\omega' \in E' : \omega'^{s-t} = \omega^{*s-t}\})}{p(\{\omega^{*s-t}\})} \text{ and } p_i(F' \mid s-t, \omega) = \frac{p(\{\omega' \in F' : \omega'^{s-t} = \omega^{*s-t}\})}{p(\{\omega^{*s-t}\})}.$$

Therefore,

$$\frac{p_i(E' \mid t, \omega)}{p_i(F' \mid t, \omega)} = \frac{p_i(E')}{p_i(F')}.$$

Since $\mathcal{S}^{s-t, s-t+1, \ldots}$ is homeomorphic to Ω, $p_i(\cdot \mid s-t, \omega)$ can be transformed into a probability in $\mathscr{P}$. Let $\mathbf{p}' \in \mathscr{P}^n$ be such that $p'_i(\omega) = p_i(\omega^{*s-t}\omega \mid s-t, \omega^*)$ for each i. Let $E'' = (\omega^*_{s-t+1}, \ldots, \omega^*_s)$ and $F'' = (\omega^*_{s-t+1}, \ldots, \omega^*_{s-1}, \omega_s)$. We therefore have $p'_i(E'') = p_i(E' \mid s-t, \omega^*)$ and $p'_i(F'') = p_i(F' \mid s-t, \omega^*)$. Hence, for each i,

$$\frac{p'_i(E'')}{p_i(F'')} = \frac{p_i(E)}{p_i(F)}.$$

Since $E, F, E'', F'' \in \mathcal{S}^t$, we have

$$\varphi^t\left(\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)}\right) = \varphi^t\left(\frac{p'_1(E'')}{p_1(F'')}, \ldots, \frac{p'_n(E'')}{p_n(F'')}\right),$$

which implies

$$\frac{k(\mathbf{p})(E)}{k(\mathbf{p})(F)} = \frac{k(\mathbf{p}')(E'')}{k(\mathbf{p}')(F'')}.$$

According to DC,

$$k(\mathbf{p}')(E'') = k(\mathbf{p})(E' \mid t, \omega^*) \text{ and } k(\mathbf{p}')(F'') = k(\mathbf{p})(F' \mid t, \omega^*).$$

That is,

$$\frac{k(\mathbf{p}')(E'')}{k(\mathbf{p}')(F'')} = \frac{k(\mathbf{p})(E' \mid t, \omega^*)}{k(\mathbf{p})(F' \mid t, \omega^*)} = \frac{k(\mathbf{p})(E')}{k(\mathbf{p})(F')}$$

Since $\varphi^t(\mathbf{a}) = \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(F)}$ and $\varphi^s(\mathbf{a}) = \frac{k(\mathbf{p})(E')}{k(\mathbf{p})(F')}$, we must have $\varphi^t(\mathbf{a}) = \varphi^s(\mathbf{a})$.

□

Based on the above result, the value of function φ^t is independent of time t. Therefore, we can subtract the superscript t and use function φ to represent φ^t.

Lemma B10. *For any* $\mathbf{a}, \mathbf{a}' \in (0, \infty)^n$, $\varphi(\mathbf{a} \cdot \mathbf{a}') = \varphi(\mathbf{a}) \cdot \varphi(\mathbf{a}')$.

Proof. Pick any $\mathbf{a}, \mathbf{a}' \in (0, \infty)^n$. There exist $E, F, G \in \mathcal{S}^t$ and $\mathbf{p} \in \mathscr{P}^n$ such that $\mathbf{a} = (\frac{p_1(E)}{p_1(F)}, \ldots, \frac{p_n(E)}{p_n(F)})$ and $\mathbf{a}' = (\frac{p_1(F)}{p_1(G)}, \ldots, \frac{p_n(F)}{p_n(G)})$. Therefore, $\mathbf{a} \cdot \mathbf{a}' = (\frac{p_1(E)}{p_1(G)}, \ldots, \frac{p_n(E)}{p_n(G)})$. By definition of φ, we have

$$\begin{aligned}
\varphi(\mathbf{a} \cdot \mathbf{a}') &= \varphi(\frac{p_1(E)}{p_1(G)}, \ldots, \frac{p_n(E)}{p_n(G)}) \\
&= \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(G)} \\
&= \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(F)} \cdot \frac{k(\mathbf{p})(F)}{k(\mathbf{p})(G)} \\
&= \varphi(\mathbf{a}) \cdot \varphi(\mathbf{a}').
\end{aligned}$$

□

Lemma B11. *φ is continuous, increasing and homogeneous of degree one.*

Proof. Let $\mathbf{a}, \mathbf{a}' \in (0, \infty)^n$ be such that $\mathbf{a} \geq \mathbf{a}'$. Accordingly, pick $\mathbf{p} \in \mathscr{P}^n$ and $E, F, G \in \mathcal{S}^t$ such that $\mathbf{a} = (\frac{p_1(E)}{p_1(G)}, \ldots, \frac{p_n(E)}{p_n(G)})$ and $\mathbf{a}' = (\frac{p_1(F)}{p_1(G)}, \ldots, \frac{p_n(F)}{p_n(G)})$. By assumption, for all $i \in \mathcal{I}$, $p_i(E) \geq p_i(F)$. Hence, $p(E) \geq p(F)$. Therefore, $k(\mathbf{p})(E) \geq k(\mathbf{p})(F)$ by SU (ii). According to the definition of

φ, we have

$$\begin{aligned}\frac{\varphi(\mathbf{a})}{\varphi(\mathbf{a}')} &= \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(G)} \cdot \frac{k(\mathbf{p})(G)}{k(\mathbf{p})(F)} \\ &= \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(F)} \\ &\geq 1.\end{aligned}$$

To see the homogeneous of degree one, first let $\mathbf{a}, \mathbf{a}' \in (0,\infty)^n$ and $\alpha \in (0,1]$ be such that $\mathbf{a} = \alpha \cdot \mathbf{a}'$. Accordingly, pick $\mathbf{p} \in \mathscr{P}^n$ and $E, F, G \in \mathcal{S}^t$ such that $\mathbf{a} = (\frac{p_1(E)}{p_1(G)}, \ldots, \frac{p_n(E)}{p_n(G)})$ and $\mathbf{a}' = (\frac{p_1(F)}{p_1(G)}, \ldots, \frac{p_n(F)}{p_n(G)})$. So, $p_i(F) = \alpha p_i(E)$ for each i. Therefore, IIA (iii) implies that $k(\mathbf{p})(E) = \alpha \cdot k(\mathbf{p})(F)$. Hence,

$$\begin{aligned}\varphi(\mathbf{a}) &= \frac{k(\mathbf{p})(E)}{k(\mathbf{p})(G)} \\ &= \alpha \cdot \frac{k(\mathbf{p})(F)}{k(\mathbf{p})(G)} \\ &= \alpha \cdot \varphi(\mathbf{a}').\end{aligned}$$

Similarly, for the case where $\mathbf{a} = \alpha \cdot \mathbf{a}'$ and $\alpha > 1$. It's clear that $\mathbf{a}' = \frac{1}{\alpha}\mathbf{a}$, where $\frac{1}{\alpha} \in (0,1)$. Applying the above process, we have $\varphi(\mathbf{a}') = \frac{1}{\alpha}\varphi(\mathbf{a})$, which is $\varphi(\mathbf{a}) = \alpha \cdot \varphi(\mathbf{a}')$. Hence, φ is homogeneous of degree one. Finally, continuity follows directly from the axiom of Continuity. □

Lemma B12. *There exist* $\gamma_1, \ldots, \gamma_n \in \mathbb{R}_+$ *with* $\sum \gamma_i = 1$ *such that* $\varphi(\mathbf{a}) = \prod_{i=1}^n a_i^{\gamma_i}$.

Proof. Applying Cauchy's equation again, the function φ must have the following form: there exist $\gamma_1, \ldots, \gamma_n$ such that for all $\mathbf{a} \in (0,\infty)^n$,

$$\varphi(\mathbf{a}) = \prod_{i=1}^n a_i^{\gamma_i}.$$

By homogeneous of degree one, for any $\alpha \in (0,\infty)$, we have $\varphi(\alpha \cdot \mathbb{1}) = \alpha$. This implies that

$$\varphi(\alpha \cdot \mathbb{1}) = \prod_{i=1}^n \alpha^{\gamma_i} = \alpha^{\sum_{i=1}^n \gamma_i} = \alpha.$$

Therefore, we must have $\sum_{i=1}^n \gamma_i = 1$. That each γ_i is non-negative follows directly from the fact that φ is increasing. □

Now, we prove the necessity of the theorem. In fact, it is trivial to show that Continuity, US, SU and IIA (i)&(ii) are necessary. We only demonstrate that DC and IIA (iii) are necessary.

Necessity of Dynamic Consistency: Assume that, for $\{W_{t,\omega}\}$ and $\{\mathbf{W}_{t,\omega}\}$, $f(\mathbf{W}_0) = W_0$. Let W_i be equipped with (u_i, d_i, p_i) for each $0 \le i \le n$. Therefore, we have, for some nonnegative numbers $\{\alpha_i\},\{\beta_i\},\{\gamma_i\}$

$$u_0(x) = \sum_{i=1}^{n} \alpha_i u_i(x), \text{ and } d_0(\tau) = \prod_{i=1}^{n} [d_i(\tau)]^{\beta_i}, \text{ and } p_0(\omega^t) \propto \prod_{i=1}^{n} [p_i(\omega^t)]^{\gamma_i}, \forall \omega^t \in \mathscr{S}.$$

Accordingly, for (t,ω),

$$\begin{aligned} d_{0,t}(\tau) &= \frac{d_0(t+\tau)}{d_0(t)} \\ &= \frac{\prod_{i=1}^{n} [d_i(t+\tau)]^{\beta_i}}{\prod_{i=1}^{n} [d_i(t)]^{\beta_i}} \\ &= \prod_{i=1}^{n} \left[\frac{d_i(t+\tau)}{d_i(t)}\right]^{\beta_i} \\ &= \prod_{i=1}^{n} \left[d_{i,t}(\tau)\right]^{\beta_i} \end{aligned}$$

Also, for $\hat{\omega}^\tau$ where $\tau > t$ and $\hat{\omega}^t = \omega^t$,

$$\begin{aligned} p_{0,(t,\omega)}(\{\hat{\omega}^\tau\}) &= \frac{p_0(\{\omega' : \omega'^\tau = \hat{\omega}^\tau\})}{p_0(\omega^t)} \\ &= \frac{\prod_{i=1}^{n} [p_i(\{\omega' : \omega'^\tau = \hat{\omega}^\tau\})]^{\gamma_i}}{\prod_{i=1}^{n} [p_i(\omega^t)]^{\gamma_i}} \\ &= \prod_{i=1}^{n} \left[\frac{p_i(\{\omega' : \omega'^\tau = \hat{\omega}^\tau\})}{p_i(\omega^t)}\right]^{\gamma_i} \\ &= \prod_{i=1}^{n} \left[p_{i,(t,\omega)}(\hat{\omega}^\tau)\right]^{\gamma_i} \end{aligned}$$

Clearly, for $\hat{\omega}^\tau$ not satisfying $\tau > t$ and $\hat{\omega}^t = \omega^t$, we know that $p_{i,(t,\omega)}(\hat{\omega}^\tau) = 0$ for each $0 \le i \le n$.

Necessity of IIA(iii): Assume that, for $\mathbf{p}, \mathbf{p}'$, each p_i dominates p_i' in probability-ratio between ω^t and ω'^t. By definition of dominance in probability-ratio, we know that for $1 \leq i \leq n$,

$$\frac{p_i(\omega^t)}{p_i(\omega'^t)} \geq \frac{p'_i(\omega^t)}{p'_i(\omega'^t)}.$$

Therefore,

$$\prod_{i=1}^{n}\Big[\frac{p_i(\omega^t)}{p_i(\omega'^t)}\Big]^{\gamma_i} \geq \prod_{i=1}^{n}\Big[\frac{p'_i(\omega^t)}{p'_i(\omega'^t)}\Big]^{\gamma_i}$$

According to the geometric probability aggregation rule, we have $\frac{k(\mathbf{p})(\omega^t)}{k(\mathbf{p})(\omega'^t)} \geq \frac{k(\mathbf{p}')(\omega^t)}{k(\mathbf{p}')(\omega'^t)}$.

C Proof of Theorem 2

In fact, we only need to prove the discounting aggregation is geometric. The proofs of other two aggregation rules are identical to the proof of Theorem 1.

Lemma C13. *For $\mathbf{d} \in \mathscr{D}^n$, if $\mathbf{d}(s) = \gamma\mathbf{d}(t)$ for some $s,t \in \mathcal{T}$ and $\gamma > 0$, then $h(\mathbf{d})(s) = \gamma h(\mathbf{d})(t)$.*

Proof. Let $\mathbf{d}$ be such that $\mathbf{d}(s) = \gamma\mathbf{d}(t)$. Wlog, assume that $\gamma < 1$. So, $t < s$ since $\mathbf{d}$ is strictly decreasing. Since preference profile follows conditional preferences, we must have $\mathbf{d}_t(0) = \mathbb{1}$ and $\mathbf{d}_t(s-t) = \gamma \cdot \mathbb{1}$.

Let consumption stream $\ell, \ell' \in X^{\mathcal{T}}$ be such that

$$\ell = (x, y, \ldots, y, \underbrace{z}_{s-t}, y, \ldots) \quad \text{and} \quad \ell' = (w, y, \ldots, y, \underbrace{x}_{s-t}, y, \ldots).$$

Clearly, ℓ, ℓ' are co-diperiodic. Take $\mathbf{W}$ equipped with $\mathbf{d}'$ and $\mathbf{u}$, where $\mathbf{d}' = \mathbf{d}_t$ and $\mathbf{u}$ be such that $\mathbf{u}(x) = \mathbb{1}$, $\mathbf{u}(y) = 0$, $\mathbf{u}(z) = 2 \cdot \mathbb{1}$ and $\mathbf{u}(w) = (1+\gamma) \cdot \mathbb{1}$. Therefore,

$$\mathbf{u} \circ \ell = [\mathbb{1}, 0, \ldots, 0, \underbrace{2 \cdot \mathbb{1}}_{s-t}, 0, \ldots] \text{ and } \mathbf{u} \circ \ell' = [(1+\gamma) \cdot \mathbb{1}, 0, \ldots, 0, \underbrace{\mathbb{1}}_{s-t}, 0, \ldots]$$

Hence, $\mathbf{W}(\ell) = (1+2\gamma) \cdot \mathbb{1}$ and $\mathbf{W}(\ell') = (1+2\gamma) \cdot \mathbb{1}$. By SU (iii)$^+$, we have $f(\mathbf{W})(\ell) = f(\mathbf{W})(\ell')$. According to the felicity aggregation result,

$$g(\mathbf{u})(x) = 1 \quad \text{and} \quad g(\mathbf{u})(y) = 0,$$
$$g(\mathbf{u})(z) = 2 \quad \text{and} \quad g(\mathbf{u})(w) = 1 + \gamma.$$

Also, $f(\mathbf{W})$ is equipped with $h(\mathbf{d}')$. Thus,

$$f(\mathbf{W})(\ell) = 1 + 2 \cdot h(\mathbf{d}')(t-s) \quad \text{and} \quad f(\mathbf{W})(\ell') = 1 + \gamma + h(\mathbf{d}')(t-s).$$

Hence, $h(\mathbf{d}')(t-s) = \gamma$. Since $\mathbf{d}_t = \mathbf{d}'$, it implies $h(\mathbf{d}') = h(\mathbf{d}_t) = \gamma$. DC requires that $h(\mathbf{d})(s) = \gamma h(\mathbf{d})(t)$.

□

Lemma C14. *There is a continuous and increasing function $\varphi : (0,\infty)^n \to (0,\infty)$ such that, for all s,t in $\mathcal{T}$,*

$$\varphi\Big(\frac{d_1(s)}{d_1(t)},\ldots,\frac{d_n(s)}{d_n(t)}\Big) = \frac{h(\mathbf{d})(s)}{h(\mathbf{d})(t)}$$

for all $\mathbf{d}$.

Proof. Notice that

$$\Big\{\mathbf{a} = \Big(\frac{d_1(s)}{d_1(t)},\ldots,\frac{d_n(s)}{d_n(t)}\Big) : \mathbf{d} \in \mathscr{D} \text{ and } s,t \in \mathcal{T}\Big\} = (0,\infty)^n.$$

We first consider the case where $\mathbf{a} \in (0,1]^\infty$. Then, there exist $\mathbf{d}$ and $s \geq t$ such that $\mathbf{a} = \Big(\frac{d_1(s)}{d_1(t)},\ldots,\frac{d_n(s)}{d_n(t)}\Big)$. Let $\mathbf{d}, \hat{\mathbf{d}}$ be such that for some $s,t,\hat{s},\hat{t}$

$$\mathbf{a} = \Big(\frac{d^1(s)}{d^1(t)},\ldots,\frac{d^n(s)}{d^n(t)}\Big) = \Big(\frac{\hat{d}^1(\hat{s})}{\hat{d}^1(\hat{t})},\ldots,\frac{\hat{d}^n(\hat{s})}{\hat{d}^n(\hat{t})}\Big).$$

To demonstrate the function φ is well-defined, we want to show that for all $\mathbf{u}$,

$$\frac{h(\mathbf{d})(s)}{h(\mathbf{d})(t)} = \frac{h(\hat{\mathbf{d}})(\hat{s})}{h(\hat{\mathbf{d}})(\hat{t})}.$$

Note that $\mathbf{d}_t(s-t) = \hat{\mathbf{d}}_{\hat{t}}(\hat{s}-\hat{t})$. Let $\mathbf{u}$ be such that, for $x,y,z \in X$, $\mathbf{u}(x) = \mathbb{1}$, $\mathbf{u}(y) = 0$ and $\mathbf{u}(z) = \mathbf{d}_t(s-t)$. Let $\ell,\ell',\ell'' \in X^\infty$ be such that

$$\ell = (x,y,\ldots,y,\underbrace{x}_{s-t},y,\ldots) \text{ and } \ell' = (x,y,\ldots,y,\underbrace{x}_{\hat{s}-\hat{t}},y,\ldots) \text{ and } \ell'' = (z,y,y,\ldots).$$

Let $\mathbf{W}'$ and $\mathbf{W}''$ be, respectively, equipped with $(\mathbf{d}' = \mathbf{d}_t, \mathbf{u})$ and $(\mathbf{d}'' = \mathbf{d}_{\hat{t}}, \mathbf{u})$. Since ℓ,ℓ'' are co-diperiodic and $\mathbf{W}'(\ell) = \mathbf{W}'(\ell'') = \mathbf{d}_t(s-t)$, SU (iii)$^+$ implies that $f(\mathbf{W}')(\ell) = f(\mathbf{W}')(\ell'')$.

Since $g(\mathbf{u})(x) = 1$ and $g(\mathbf{u})(y) = 0$, we know

$$f(\mathbf{W}')(\ell) = 1 + h(\mathbf{d}')(s-t) \quad \text{and} \quad f(\mathbf{W}')(\ell'') = g(\mathbf{u})(z).$$

Similarly, ℓ', ℓ'' are co-diperiodic and $\mathbf{W}''(\ell') = \mathbf{W}''(\ell'')$, which, by SU (iii)$^+$ implies that $f(\mathbf{W}'')(\ell') = f(\mathbf{W}'')(\ell'')$. And we have,

$$f(\mathbf{W}'')(\ell') = 1 + h(\mathbf{d}'')(\hat{s}-\hat{t}) \quad \text{and} \quad f(\mathbf{W}'')(\ell'') = g(\mathbf{u})(z).$$

Consequently,

$$h(\mathbf{d}')(s-t) = h(\mathbf{d}'')(\hat{s}-\hat{t}),$$

which implies $h(\mathbf{d}_t)(s-t) = h(\mathbf{d}_{\hat{t}})(\hat{s}-\hat{t})$. Hence, DC implies

$$\frac{h(\mathbf{d})(s)}{h(\mathbf{d})(t)} = \frac{h(\hat{\mathbf{d}})(\hat{s})}{h(\hat{\mathbf{d}})(\hat{t})}.$$

The proof of the other case where $\mathbf{a} \in [1,\infty)^n$ is similar simply because $\frac{1}{\mathbf{a}} \in (0,1]^n$. Also, the continuity of φ follows directly from the axiom of Continuity.

To see φ is increasing, take $\mathbf{a}, \mathbf{b} \in (0,\infty)^n$ such that $\mathbf{a} \geq \mathbf{b}$. There are three cases to consider. Suppose that $\mathbf{a}, \mathbf{b} \in (0,1]^n$. Then, there exist $\mathbf{d}$ and $r \geq s \geq t$ such that

$$\mathbf{a} = \left(\frac{d_1(s)}{d_1(t)}, \ldots, \frac{d_n(s)}{d_n(t)}\right) \quad \text{and} \quad \mathbf{b} = \left(\frac{d_1(r)}{d_1(t)}, \ldots, \frac{d_n(r)}{d_n(t)}\right).$$

Therefore,

$$\frac{\varphi(\mathbf{a})}{\varphi(\mathbf{b})} = \frac{h(\mathbf{d})(s)}{h(\mathbf{d})(t)} \cdot \frac{h(\mathbf{d})(t)}{h(\mathbf{d})(r)} = \frac{h(\mathbf{d})(s)}{h(\mathbf{d})(r)} \geq 1.$$

The case where $\mathbf{a}, \mathbf{b} \in [1,\infty)^n$ is symmetric and the proof is very similar. The third case is $\mathbf{a} \in [1,\infty)^n$ and $\mathbf{b} \in (0,1]^n$. There are $\mathbf{d}$ and $s \geq t$, $s' \geq t'$ such that

$$\mathbf{a} = \left(\frac{d_1(t)}{d_1(s)}, \ldots, \frac{d_n(t)}{d_n(s)}\right) \quad \text{and} \quad \mathbf{b} = \left(\frac{d_1(s')}{d_1(t')}, \ldots, \frac{d_n(s')}{d_n(t')}\right).$$

Therefore,

$$\varphi(\mathbf{a}) = \frac{h(\mathbf{d})(t)}{h(\mathbf{d})(s)} \geq 1 \geq \frac{h(\mathbf{d})(s')}{h(\mathbf{d})(t')} = \varphi(\mathbf{b}).$$

□

Lemma C15. *For any* $\mathbf{a}, \mathbf{b} \in (0,\infty)^n$, $\varphi(\mathbf{a} \cdot \mathbf{b}) = \varphi(\mathbf{a}) \cdot \varphi(\mathbf{b})$.

Proof. Take any $\mathbf{a}, \mathbf{b} \in (0, \infty)^n$, there exist $\mathbf{d}$ and r, s, t such that

$$\mathbf{a} = \left(\frac{d_1(r)}{d_1(s)}, \ldots, \frac{d_n(r)}{d_n(s)}\right) \text{ and } \mathbf{b} = \left(\frac{d_1(s)}{d_1(t)}, \ldots, \frac{d_n(s)}{d_n(t)}\right).$$

By definition of φ, it is immediate that

$$\varphi(\mathbf{a} \cdot \mathbf{b}) = \varphi\left(\frac{d_1(r)}{d_1(t)}, \ldots, \frac{d_n(r)}{d_n(t)}\right) = \frac{h(\mathbf{d})(r)}{h(\mathbf{d})(t)} = \frac{h(\mathbf{d})(r)}{h(\mathbf{d})(s)} \cdot \frac{h(\mathbf{d})(s)}{h(\mathbf{d})(t)} = \varphi(\mathbf{a}) \cdot \varphi(\mathbf{b}).$$

□

Lemma C16. *There exist non-negative numbers $\beta_1, \ldots, \beta_n$ with $\sum_{i=1}^n \beta_i = 1$ such that, for any $(\mathbf{u}, \mathbf{d})$, $h(\mathbf{d}) = \prod_{i=1}^n d_i^{\beta_i}$.*

Proof. Define function $g : \mathbb{R}^n \to \mathbb{R}$ by, for all $\mathbf{a} \in \mathbb{R}^n$,

$$g(\mathbf{a}) = \ln(\varphi(\exp(\mathbf{a}))).$$

Clearly, for all $\mathbf{a}, \mathbf{b}$,

$$g(\mathbf{a}) + g(\mathbf{b}) = \ln(\varphi(\exp(\mathbf{a}))) + \ln(\varphi(\exp(\mathbf{b}))) = \ln(\varphi(\exp(\mathbf{a} + \mathbf{b}))) = g(\mathbf{a} + \mathbf{b}).$$

Since φ is continuous, we know that $\varphi(b \cdot \mathbb{1}) = b$ for any $b \in (0, \infty)$. So, take $a \in \mathbb{R}$,

$$g(a \cdot \mathbb{1}) = \ln(\varphi(\exp(a \cdot \mathbb{1}))) = a.$$

Furthermore, continuity of φ implies g is also continuous. Hence, applying Cauchy's equation, the function g must have the following form:

$$g(\mathbf{a}) = \beta_1 a_1 + \cdots + \beta_n a_n.$$

Also, $g(a \cdot \mathbb{1}) = a$ implies that $\sum_{i=1}^n \beta_i = 1$. Since φ is increasing, function g is also increasing, which implies each β_i is non-negative. Now, take any $\mathbf{a} \in (0, \infty)^n$,

$$\ln(\varphi(\mathbf{a})) = g(\ln(\mathbf{a})) = \sum_{i=1}^n \beta_i \ln(a_i).$$

Therefore, $\varphi(\mathbf{a}) = \prod_{i=1}^n a_i^{\beta_i}$.

Take any $\mathbf{d}$ and $t \in \mathcal{T}$. Clearly $\mathbf{d}(t) = \frac{\mathbf{d}(t)}{\mathbf{d}(1)} \in (0, \infty)^n$. Hence,

$$h(\mathbf{d})(t) = \varphi(\mathbf{d}(t)) = \prod_{i=1}^{n} [d_i(t)]^{\beta_i} \quad \text{for all } t \in \mathcal{T}.$$

□

References

Alon, Shiri, and Gabi Gayer. 2016. "Utilitarian Preferences With Multiple Priors." *Econometrica*, 84(3): 1181–1201.

Bernheim, B. Douglas. 1989. "Intergenerational Altruism, Dynastic Equilibria and Social Welfare." *The Review of Economic Studies*, 56(1): 119–128.

Bewley, Truman F. 2002. "Knightian decision theory. Part I." *Decisions in Economics and Finance*, 25(2): 79–110.

Billot, Antoine, and Xiangyu Qu. 2021*a*. "Utilitarian Aggregation with Heterogeneous Beliefs." *American Economic Journal: Microeconomics*, 13(3): 112–23.

Billot, Antoine, and Xiangyu Qu. 2021*b*. "Perfect Altruism Breeds Time Consistency." working paper or preprint.

Buchanan, J.M., and G. Tullock. 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy. Ann Arbor paperbacks*, University of Michigan Press.

Caplin, Andrew, and John Leahy. 2004. "The Social Discount Rate." *Journal of Political Economy*, 112(6): 1257–1268.

Chambers, Christopher P., and Federico Echenique. 2018. "On Multiple Discount Rates." *Econometrica*, 86(4): 1325–1346.

Chambers, Christopher P., and Takashi Hayashi. 2006. "Preference aggregation under uncertainty: Savage vs. Pareto." *Games and Economic Behavior*, 54(2): 430–440.

Chandrasekher, Madhav, Mira Frick, Ryota Iijima, and Yves Le Yaouanq. 2022. "Dual-Self Representations of Ambiguity Preferences." *Econometrica*, 90(3): 1029–1061.

Chichilnisky, Graciela, Peter J. Hammond, and Nicholas Stern. 2020. "Fundamental utilitarianism and intergenerational equity with extinction discounting." *Social Choice and Welfare*, 54(2): 397–427.

Christensen, Jens Peter Reus. 1972. "On sets of Haar measure zero in abelian Polish groups." *Israel Journal of Mathematics*, 13(3): 255–260.

Clarke, Frank H. 1990. *Optimization and Nonsmooth Analysis*. Society for Industrial and Applied Mathematics.

Crès, Hervé, Itzhak Gilboa, and Nicolas Vieille. 2011. "Aggregation of multiple prior opinions." *Journal of Economic Theory*, 146(6): 2563–2582.

Danan, Eric, Thibault Gajdos, Brian Hill, and Jean-Marc Tallon. 2016. "Robust Social Decisions." *American Economic Review*, 106(9): 2407–25.

Dietrich, Franz. 2021. "Fully Bayesian aggregation." *Journal of Economic Theory*, 194: 105255.

Drouhin, Nicolas. 2020. "Non-stationary additive utility and time consistency." *Journal of Mathematical Economics*, 86: 1 – 14.

Drugeon, Jean-Pierre, and Bertrand Wigniolle. 2020. "On Markovian collective choice with heterogeneous quasi-hyperbolic discounting." *Economic Theory*.

Drupp, Moritz A., Mark C. Freeman, Ben Groom, and Frikk Nesje. 2018. "Discounting Disentangled." *American Economic Journal: Economic Policy*, 10(4): 109–34.

d'Aspremont, Claude. 1985. "Axioms for social welfare orderings." *Social goals and social organization: Essays in memory of Elisha Pazner*, 19–76.

Ellsberg, Daniel. 1961. "Risk, Ambiguity, and the Savage Axioms." *The Quarterly Journal of Economics*, 75(4): 643–669.

Epstein, Larry G., and Martin Schneider. 2003. "Recursive multiple-priors." *Journal of Economic Theory*, 113(1): 1–31.

Farhi, Emmanuel, and Iván Werning. 2007. "Inequality and Social Discounting." *Journal of Political Economy*, 115(3): 365–402.

Feldstein, M. S. 1964. "The Social Time Preference Discount Rate in Cost Benefit Analysis." *The Economic Journal*, 74(294): 360–379.

Feng, Tangren, and Shaowei Ke. 2018. "Social Discounting and Intergenerational Pareto." *Econometrica*, 86(5): 1537–1567.

Galaabaatar, Tsogbadral, and Edi Karni. 2013. "Subjective Expected Utility With Incomplete Preferences." *Econometrica*, 81(1): 255–284.

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci. 2004. "Differentiating ambiguity and ambiguity attitude." *Journal of Economic Theory*, 118(2): 133–173.

Gilboa, Itzhak, and David Schmeidler. 1989. "Maxmin expected utility with non-unique prior." *Journal of Mathematical Economics*, 18(2): 141 – 153.

Gilboa, Itzhak, Dov Samet, and David Schmeidler. 2004. "Utilitarian Aggregation of Beliefs and Tastes." *Journal of Political Economy*, 112(4): 932–938.

Halevy, Yoram. 2015. "Time Consistency: Stationarity and Time Invariance: Time Consistency." *Econometrica*, 83(1): 335–352.

Harsanyi, John C. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy*, 63(4): 309–321.

Heal, Geoffrey, and Antony Millner. 2018. "Chapter 10 - Uncertainty and ambiguity in environmental economics: conceptual issues." In *Handbook of Environmental Economics*. Vol. 4 of *Handbook of Environmental Economics*, , ed. Partha Dasgupta, Subhrendu K. Pattanayak and V. Kerry Smith, 439–468. Elsevier.

Herstein, I. N., and John Milnor. 1953. "An Axiomatic Approach to Measurable Utility." *Econometrica*, 21(2): 291–297.

Hsu, Ming, Meghana Bhatt, Ralph Adolphs, Daniel Tranel, and Colin F Camerer. 2005. "Neural systems responding to degrees of uncertainty in human decision-making." *Science*, 310(5754): 1680–1683.

Hylland, Aanund, and Richard Zeckhauser. 1979. "The Impossibility of Bayesian Group Decision Making with Separate Aggregation of Beliefs and Values." *Econometrica*, 47(6): 1321–1336.

Jackson, Matthew O., and Leeat Yariv. 2014. "Present Bias and Collective Dynamic Choice in the Lab." *American Economic Review*, 104(12): 4184–4204.

Jackson, Matthew O., and Leeat Yariv. 2015. "Collective Dynamic Choice: The Necessity of Time Inconsistency." *American Economic Journal: Microeconomics*, 7(4): 150–178.

Johnsen, Thore H., and John B. Donaldson. 1985. "The Structure of Intertemporal Preferences under Uncertainty and Time Consistent Plans." *Econometrica*, 53(6): 1451–1458.

Koopmans, Tjalling C. 1960. "Stationary Ordinal Utility and Impatience." *Econometrica*, 28(2): 287.

Marglin, Stephen A. 1963. "The Social Rate of Discount and The Optimal Rate of Investment." *The Quarterly Journal of Economics*, 77(1): 95–111.

Millner, Antony, and Geoffrey Heal. 2018*a*. "Time consistency and time invariance in collective intertemporal choice." *Journal of Economic Theory*, 176: 158–169.

Millner, Antony, and Geoffrey Heal. 2018*b*. "Discounting by Committee." *Journal of Public Economics*, 167: 91–104.

Mongin, Philippe. 1994. "Harsanyi's Aggregation Theorem: multi-profile version and unsettled questions." *Social Choice and Welfare*, 11.

Mongin, Philippe. 1995. "Consistent Bayesian Aggregation." *Journal of Economic Theory*, 66(2): 313 – 351.

Pivato, Marcus. 2022. "Bayesian social aggregation with accumulating evidence." *Journal of Economic Theory*, 200: 105399.

Qu, Xiangyu. 2017. "Separate aggregation of beliefs and values under ambiguity." *Economic Theory*, 63(2): 503–519.

Savage, Leonard. 1954. *The foundations of statistics.* (Second Edition in 1972), New York: Dover.

Schmeidler, David. 1989. "Subjective Probability and Expected Utility without Additivity." *Econometrica*, 57(3): 571–587.

Weitzman, Martin L. 2001. "Gamma Discounting." *American Economic Review*, 91(1): 260–271.

Zuber, Stéphane. 2011. "The aggregation of preferences: can we ignore the past?" *Theory and Decision*, 70(3): 367–384.