Intergenerational altruism and time consistency*

Łukasz Balbus† Kevin Reffett‡ Łukasz Woźny§

May 2021

Abstract

We consider a class of infinite horizon, stochastic, non-stationary dynastic consumption-savings models with a general forms of recursive, time-varying altruistic preferences including direct and indirect pure altruism as well as paternalistic altruism. It is well-known such models lead to time-inconsistent dynastic preferences. Within this class of economies, we propose a novel set-iterative procedure for characterizing all Markov perfect time-consistent solutions in the space of increasing investments. Our approach involves both: value functions *and* policy iterations. We prove existence of Markov Perfect equilibria in stationary, periodic and also non-stationary strategies. We provide numerous applications to altruistic growth models, behavioral discounting models and collective household models and also discuss the role of various certainty equivalence operators.
Keywords: Nonpaternalistic altruism, Paternalistic altruism, Time consistency; Existence; Markovian self-generation; Markov equilibrium; Periodic-equilibria; Approximation.
JEL classification: C61, C73

1 Introduction

The paper proposes a new set-iterative approach for characterizing all Markov perfect time-consistent equilibria in a class of dynastic consumption-savings models with recursive altruistic preferences. Building on the recent axiomatic work

* The project was financed by the National Science Center: NCN grant number UMO-2019/35/B/HS4/00346.
† Faculty of Mathematics, Computer Sciences and Econometrics, University of Zielona Góra, Poland.
‡ Department of Economics, Arizona State University, USA.
§ Department of Quantitative Economics, Warsaw School of Economics, Warsaw, Poland.

on direct pure altruism of Galperti and Strulovici (2017), as well as motivated by a long line of the related work on nonpaternalistic and paternalistic altruism over the last four decades,[1] we consider a dynastic choice problem of generations connected by recursive preferences for both the case of pure (direct or indirect) as well as paternalistic altruism. The class of recursive preferences we consider is very general, allow for aggregators that are time-varying, non-stationary, defined in settings with uncertainty, and are in general time-inconsistent.

In such a framework, we define an appropriate intergenerational dynamic game for the study of time-consistent / subgame perfect equilibrium. Relative to this game, we propose a strategic dynamic programming approach that restricts attention to the space of *Markovian strategies*. This restriction leads to new set-iterative method that differs from the typical self-generation approach that has been proposed to solving repeated/dynamic games in the existing literature (e.g., Abreu et al. (1990)).[2] Our approach applies in the setting of a dynamic/stochastic game, and then maps into function spaces for both policies and values. In the end, for our class of models, we are able to characterize *all* the monotone Markov Perfect Equilibrium in stationary, periodic, and nonstationary strategies via the largest fixed point (under set inclusion) of our Markovian self-generation operator.

Our paper builds on the axiomatic work of Galperti and Strulovici (2017), as well as related work on direct pure altruism models in many papers in the literature including Ray (1987), Saez-Marti and Weibull (2005), Pearce (2008), Fels and Zeckhauser (2008), among other. With that said, our results lead to three significant new directions : (i) by applying a general recursive aggregator approach, for the direct pure altruistic preference case, we are able to prove the existence of time consistent equilibria in dynastic choice models for general *time-varying* dynastic

[1] For example, among many papers in this extensive literature, for nonpaternalistic altruism, see Barro (1974), Ray (1987), Saez-Marti and Weibull (2005), Pearce (2008); for paternalistic altruism, see Koopmans (1960), Bernheim and Ray (1986), Leininger (1986), Asheim (2010).

[2] For the recent literature on correspondence-based strategic dynamic programming, as well as an explanation of the traditional strategic dynamic programming approach to dynamic/stochastic games, see the papers of Bernheim et al. (1999), Phelan and Stacchetti (2001), Chade et al. (2008), Bernheim et al. (2015), Sleet and Ş. Yeltekin (2016), Yeltekin et al. (2017), Abreu et al. (2020a,b), among others.

preference aggregators $\{V_t\}$ without imposing conditions such as intergenerational separability or altruistic stationarity, (ii) with our single unified framework, we are also able to prove the existence of time-consistent equilibria for models of indirect pure altruism (as in Koopmans (1960), Barro (1974), Loury (1981)) and Barczyk and Kredler (2020)) as well as those with paternalistic altruism (such as Bernheim and Ray (1986), Leininger (1986), Asheim (2010), and Mookherjee and Napel (2019)), (iii) we are able to consider time-consistent dynastic choice under uncertainty, allowing for random state transitions and general forms of certainty equivalence operators.

Our approach to equilibrium existence and equilibrium construction unifies the construction of (nonstationary) Markov perfect time consistent equilibria with the typical case studied in the existing literature on Stationary Markov Perfect equilibrium, as well as provides the first sufficient conditions in the literature on *periodic* Markov Perfect Equilibria. In addition, as many dynamic models of time-inconsistent choice with behavioral discounting are special cases of direct pure altruistic models (e.g., see Galperti and Strulovici (2017), section 4), we are also able to relate our Markovian strategic dynamic programming approach to the extensive literature on time consistent equilibria in dynamic models with behavioral discounting, such as quasi-hyperbolic discounting, hyperbolic discounting and generalized behavioral discounting with a single, infinitely-lived consumer, who is given an intergenerational interpretation.

The rest of the paper is organized as follows. In the following section (2) we present the historical perspective on development of altruistic preferences literature as well as discuss our results in the context of the related papers. Section 3 presents the model, notation and our main result. Section 4 extends our results to a special case of periodic models and proves existence of periodic equilibria; in particular stationary ones. Section 5 is devoted to a generalization towards paternalistic models. Finally, section 6 presents some special cases of our results and relations to the literature. Four technical lemmas are moved to the Appendix.

2 Relations to the literature

Work studying the role of intergenerational preferences, dynastic choice, and time consistency is vast, and has become foundational topic of interest in many important literatures in economics. In some broad sense, this modern literature that focuses on altruistic preferences, bequests, and transfers (often defined in a broad sense) starts with the seminal work of both Barro (1974) on Ricardian equivalence and Becker and Tomes (1979) on bequests and the intergenerational transmission of wealth.[3]

Importantly, and more recently, there has also been a great deal of research that studies how bequests and altruism impacts the structure of dynastic precautionary savings in the presence of risk. This line of work includes the papers of Kopczuk and Lupton (2007), Ameriks et al. (2017), and Boar (2021), as well as related work on health care spending, medicaid, and dynastic savings as found in the papers of Braun et al. (2017), De Nardi et al. (2016), and Ameriks et al. (2020). All this work emphasizes the role of intergenerational preferences in explaining the transmission of wealth across generations, as well as the relationship with social insurance as a substitute for dynastic savings. See also Abbott et al. (2019). Of course, bequest and intergenerational altruism also play a critical role in understanding the role of social security as has been argued in many papers in the macroeconomics literature (see Laitner (1988), Fuster et al. (2007), and Imrohoroglu and Zhao (2018), among many others). We model allows to analyze such risk / insurance considerations by allowing for random state transition and general form of certainty equivalence.

An important new line of work focuses on the role of parenting styles and intergenerational investments for human capital development using endogenous preferences in models with indirect pure altruism and paternalism. See the work stemming from the series of recent papers by Doepke and Zilibotti (2017), Doepke

[3] Relative to the question of bequest and capital accumulation, these early papers continued with the work of Kotlikoff and Summers (1981), Abel (1987), and Bernheim and Bagwell (1988), as well as papers with strategic bequest motives as in the work of Bernheim et al. (1985), Bernheim and Ray (1986), Leininger (1986), Perozek (1998), and with the works of Becker and Tomes (1979) on bequests, transfers, intergenerational wealth and income mobility. See also Loury (1981), De Nardi (2004) and the survey by Piketty and Zucman (2015).

and Sorrenti (2019), and Doepke et al. (2019).

Additionally, there is a growing and new literature that seeks to understand the structure of dynamic collective choice, where the issue of time inconsistency naturally arise in dynamic (dynastic) models from the presence of heterogeneous discount factors (e.g, see Jackson and Yariv (2015); Lizzeri and Yariv (2017), and Millner and Heal (2018))). Such models are also related to dynamic collective household choice such as discussed in Mazzocco (2007) and Balbus et al. (2021), where again heterogenous discount rates within the collective household naturally lead to time inconsistent household preferences. In this paper, we actually show how to apply and extend the results in this literature to settings under weaker sufficient conditions.

One other related line of work where the issue of dynastic collective time consistent choice arises includes studies of intergenerational preferences, and the design of global environment policies. Here, as in dynamic collective household applications, issues of heterogeneous discount factors creates the emergence of time inconsistent preferences. See for example the recent work of Karp (2016), Gerlagh and Liski (2017), Millner (2020), and Iverson and Karp (2020). Related issues arise when one seeks to understand the structure of recursive social welfare functions and study time consistent social choice. For such models, see for example the papers of Phelan (2006), Farhi and Werning (2007), Asheim (2010), and Feng and Ke (2018). Our paper provides sufficient conditions for the existence and construction of Markov Perfect time-consistent dynastic choices for such dynamic collective decision making problems.

Finally, as shown in Galperti and Strulovici (2017), direct pure altruistic preferences include, as a special case, the quasi-hyperbolic discounting (see Strotz (1956), Pollak (1968), Phelps and Pollak (1968), Peleg and Yaari (1973), Laibson (1997), Balbus et al. (2015b, 2018), and Cao and Werning (2018)). In this paper, even the larger class of models with generalized behavioral discounting (see e.g. Chakraborty et al. (2020), Balbus et al. (2020b), and Richter (2021)) is covered.[4] Finally, as we study such issues in a stochastic dynastic setting, our

[4] Our work could be also seen as related to an emerging literature on rationalizability, time

work also relates to work relating behavioral discounting and uncertainty such as in Akerlof (1991), Prelec and Loewenstein (1991), Dasgupta and Maskin (2005), Halevy (2008), and DeJarnette et al. (2020), among others.

From a technical and methodological perspective, our work complements the literature studying subgame perfect equilibria in consumption-savings models with altruistic preferences, as well as correspondence-based strategic dynamic programming approaches to related problems in dynamic/stochastic games. For example, Balbus et al. (2016) prove existence of stationary MPE in a related class of stationary models with non-paternalistic altruism and separable preferences. Our results are more general and the method is different. In addition, relative to our Markovian strategic dynamic programming approach, some existing work has also developed self-generation approaches that restrict attention to primarily short-memory subgame perfect strategies. Two notable papers approaching the existence problem in a related manner are Doraszelski and Escobar (2012) in the context of repeated games, and Balbus and Woźny (2016) in the context of dynamic/stochastic games restrict attention to short-memory "APS" type methods. Our approach differs substantially from the former paper as the interpersonal game our time consistency problem generates is dynamic/stochastic. Relative to the later paper, as in Balbus and Woźny (2016) we develop an correspondence-based self-generation approach using Markovian strategies and values; but here, our approach works under much weaker conditions. Technically, our self-generation approach has elements that are related to the papers on quasi-hyperbolic discounting of Bernheim et al. (1999) and Bernheim et al. (2015), but works for much more general classes of dynastic altruistic models[5], and restricts attention self-generating in function spaces for both values and Markovian strategies, we are able to prove very sharp characterizations of subgame perfect equilibria.

Finally, and importantly, our work is also methodologically similar to the approach taken to the existence of equilibrium in Balbus et al. (2017). This paper,

preference, and decreasing patience such as the papers of Adams et al. (2014), Saito (2015), Echenique et al. (2020), Dziewulski (2018), and Chambers et al. (2021).

[5] See also the related to recent papers on repeated games with recursive payoffs in Obara and Park (2017).

though, allows for a much more general class of dynastic altruistic consumption-savings models. Finally, as in the recent work on APS methods and stochastic games, our work is related to Abreu et al. (2020a,b) in the sense that like these papers, we sharpen the APS approach by including the modeling of both strategies and values in the self-generation approach. In addition, as we study periodic Markov perfect equilibrium, our construction bears a relationship to the recent work on periodicity and self-generation by Berg (2017) and Berg and Kitti (2019).

3 The model and the main result

Consider an infinite sequence of generations index by $t \in \mathbb{N} = \{1, 2, \ldots\}$. Generation t has $s_t \in S = [0, \bar{s}]$ resources[6] for its own disposal. It divides it by choosing consumption[7] $c_t \in [0, s_t]$ and investment $i_t = s_t - c_t$. The resource is renewable and the amount of resources the next period s_{t+1} is a random value whose (Borel) distribution is $q_t(\cdot|i_t)$.

We adopt a more general form of direct pure altruistic preferences as recently axiomatized by Galperti and Strulovici (2017). Formally, let a sequence of consumption $\{c_t\}$ by given and introduce the t-shift operator $c^t := (c_t, c_{t+1}, \ldots)$ as well as $c^{t,\tau} := (c_t, c_{t+1}, \ldots, c_{t+\tau})$. Then, we generate recursively a sequence of utilities $\{U_t\}$ using the following formula:

$$U_t = V_t(c_t, U^{t+1}). \tag{1}$$

This formulation is general and allows for a number of special cases. It includes, for example, the standard, stationary preferences of Koopmans (1960) in the form of *indirect* pure altruism, where $V_t(c_t, U^{t+1}) = W(c_t, U_{t+1})$, and in particular time-separable case with $W(c_t, U_{t+1}) = u(c_t) + \beta U_{t+1}$.[8] It also involves more general

[6] We consider a bounded states space and hence bounded rewards. Generalizations including unbounded states space or unbounded above reward space are possible. See Balbus et al. (2020b) for a recent application.

[7] The choice of consumption and investment from $[0, s_t]$ may seem restrictive. In fact our formulation allows for more general cases with a production function $g(s_t)$, or some borrowing bounds introduced. This can be embedded in the formulation of V_t.

[8] In the paper, we use the term "direct" pure (nonpaternalistic) altruism as in Galperti and

but time-separable cases with indirect pure altruism towards T (possibly infinitely many) consecutive generations $u(c_t)+\sum_{\tau=1}^{T}\beta_t^{\tau}U_{t+\tau}$, where β_t^{τ} is the weight placed by generation t on the utility of the $t+\tau$ generation. Interestingly, our formulation allows also for time-varying and non-exponential discounting (see Balbus et al. (2020b) for a related study). This includes quasi-hyperbolic discounting with $u(c_t)+\beta\sum_{\tau=1}^{\infty}\delta^{\tau}u(c_{t+\tau})$ as a special case (see Theorem 4 and the proof of Corollary 4 in Galperti and Strulovici (2017) for a derivation). We provide further examples in section 6.

Clearly, as shown by Galperti and Strulovici (2017), these preferences defined in equation (1) are generally time-inconsistent.[9] In what follows, therefore, we define an equilibrium concept that captures the notion of a time-consistent solution in this time-inconsistent environment, but unlike the existing literature, we do *not* restrict attention to constructing Stationary Markov Perfect Equilibria for models with primitive data that are time-invariant.[10] Let $h_t : S \mapsto S$ be a Markov policy for generation t. Formally, it is a Borel measurable function such that $h_t(s) \in [0, s]$ for any $s \in S$. For $i \in S$ and integer t, let $q_t^1(\cdot|i) = q_t(\cdot|i)$. Furthermore, for any policy h_{t+1} of $t+1$ let

$$q_t^2(\cdot|i, h_{t+1}) = \int_S q_{t+1}(\cdot|h_{t+1}(s_{t+1}))q_t(ds_{t+1}|i)$$

be a transition in 2 steps. More generally, for any such $i \in S$, $\tau > 1$ and any profile $h^{t+1,\tau-1}$ applied from $t+1$ generation to $t+\tau$, define the transition in τ steps as follows:

$$q_t^{\tau+1}(\cdot|i, h^{t+1,\tau-1}) := \int_S q_{t+\tau}(\cdot|h_{t+\tau}(s_{t+\tau}))q_t^{\tau}(ds_{t+\tau}|i, h^{t+1,\tau-2}) \text{ for } i \in S.$$

Strulovici (2017) to mean each generations preferences depend on every successor generations utility (or "well-being"). The indirect case is when the pure altruistic preferences only depend on a strict subset of successor generation utilities. Similar terminology is used for the "paternalistic" altruism case, excepting the fact that each generations preferences are defined over successor generations actual consumptions.

[9] See Galperti and Strulovici (2017), Proposition 4. So apart from the special case of Koopmans (1960) with $V_t(c_t, U^{t+1}) = u(c_t) + \beta U_{t+1}$, the altruistic preferences in this paper are time inconsistent.

[10] In the next section of the paper, we consider periodic Markov Perfect Equilibria. In this case, a special case of our environment and our construction will lead to stationary Markov Perfect Equilibria.

Let $\Delta(S)$ be the set of all probability measures on S and $\mathscr{B}(S)$ be the set of all bounded Borel measurable functions on S.

We now define the Certainty Equivalent Operator, which plays a critical role in our work. We say $\hat{\mathcal{M}}_t(\cdot, \mu)$ is a *Certainty Equivalent Operator* if for any generation t, and any $\mu \in \Delta(S)$, $\hat{\mathcal{M}}_t : \mathscr{B}(S) \times \Delta(S) \mapsto \mathbb{R}$ satisfies:

- $\hat{\mathcal{M}}_t(\cdot, \mu)$ is Borel measurable;
- $\hat{\mathcal{M}}_t(\cdot, \mu)$ is monotone, that is if $f \leq g$ for μ-almost everywhere, then

$$\hat{\mathcal{M}}_t(f, \mu) \leq \hat{\mathcal{M}}_t(g, \mu);$$

- For any constant α we have $\hat{\mathcal{M}}_t(\alpha, \mu) = \alpha$;

To simplify notation, let:

$$\mathcal{M}_{i,t}(\cdot, h^{t+1,\tau-1}) := \hat{\mathcal{M}}_t\left(\cdot, q_t^{\tau+1}(\cdot|i, h^{t+1,\tau-1})\right),$$

with $\mathcal{M}_{i,t}(\cdot) := \hat{\mathcal{M}}_t(\cdot, q_t(\cdot|i))$.

Let $\mathcal{F}$ be the set of all increasing functions from S into S. On $\mathcal{F}$, we define the equivalence relation $\sim$ in which $f_1 \sim f_2$ if and only if $f_1(s) = f_2(s)$ for any s such that f_2 is continuous at s. Let $\mathscr{F}$ be the set of all equivalence classes of elements in $\mathcal{F}$ and

$$\mathscr{I} := \{h \in \mathscr{F} : h(s) \in [0, s] \text{ for all } s \in S\}.$$

Under some continuity assumptions (to be introduced in the moment), with $(f, h) \in (\mathscr{F} \times \mathscr{I})^\infty := \mathscr{V}$ we can now define the following operator for each date $t \in \mathbb{N}$ and for each $s \in S$:

$$T_t(f,h)(s) = \max_{i \in [0,s]} V_t\left(s - i, \mathcal{M}_{i,t}(f_2), \mathcal{M}_{i,t}(f_3, h^{2,0}), \ldots, \mathcal{M}_{i,t}(f_{\tau+1}, h^{2,\tau-2}), \ldots\right)$$

Similarly, define the best reply mapping:

$$H_t(f,h)(s) = \arg\max_{i \in [0,s]} V_t\left(s - i, \mathcal{M}_{i,t}(f_2), \mathcal{M}_{i,t}(f_3, h^{2,0}), \ldots, \mathcal{M}_{i,t}(f_{\tau+1}, h^{2,\tau-2}), \ldots\right). \tag{2}$$

The aforementioned problem of time-consistency of preferences in (1) is now evident from the above formulation. Indeed, whenever generation t cares about more generations than just the *immediate* descendant, one has to use policies h^2 to evaluate future streams of utilities f^3 in V_t. This may be surprising, as the model is paternalistic and hence preferences V_t do not depend directly on the consecutive generations' consumption choices. Nevertheless, we need sequence of (future) values *and policies* to evaluate current utility.

With all this in place, we can now define our equilibrium concept:

Definition 1. (Markov Perfect Equilibrium) A sequence of measurable policies $\{h_t\}$ is a Markov perfect equilibrium (or a time-consistent solution) whenever there exists a sequence of integrable values $\{f_t\}$ such that for any generation t and state $s \in S$ we have:

$$f_t(s) = T_t(f^{t+1}, h^{t+1})(s),$$
$$h_t(s) \in H_t(f^{t+1}, h^{t+1})(s).$$

Markov perfection is our main solution concept. It requires that h_t is a best response for the sequence of future generation utilities f^{t+1} each evaluated by the certainty equivalent operator using h^{t+1}. Observe, this definition precisely captures the notion of time consistency in our framework (see Strotz (1956)). It is still a rather simple equilibrium concept as Markovian strategies do not allow generations to condition future actions on states or actions of past generations. Our solution concept is also consistent with our main assumption regarding transition q_t, namely its non-atomic structure (see assumption 2 that we introduce in the moment). It implies, for example, that information about the *current* state cannot be used to *recall* past actions or past states (hence, our attention to Markovian strategies is justified). Finally, we should mention the induced equilibrium behavior can still be very rich due to nonstationarity of our environment.

Now, let us introduce more concise notation for construction of the set of Markov Perfect Equilibria. Let

$$\Phi_t(f, h)(s) := (T_t(f, h)(s), H_t(f, h)(s)).$$

More formally, note that $\Phi_t(f,h)(s) = \{T_t(f,h)\} \times [H_t(f,h)]$, where $[H_t(f,h)]$ is the set of equivalence classes whose elements are equivalent to one of the selections of $H_t(f,h)$, i.e.

$$[H_t(f,h)] := \{h \in \mathscr{I} : h \sim \gamma \text{ for some } \gamma \in Sel(H_t(f,h))\},$$

where $Sel(H_t(f,h)) := \{h : S \mapsto S : h(s) \in H_t(f,h)(s) \text{ for all } s \in S\}$.

There is a well-known natural homeomorphism between $\mathscr{F}$ and $\Delta(S)$. Given this, we endow $\mathscr{F}$ with it's inherited *weak topology*, so the convergence $\Rightarrow$ is defined as follows

$$f_n \Rightarrow f \Leftrightarrow \lim_{n\to\infty} f_n(s) = f(s), \text{ for any } s \in C_f$$

where C_f is the set of continuity points of f and $\mathscr{V}$ with the *standard product topology*. The convergence on $\mathscr{V}$ is denoted by $\Rightarrow$ as well, as is in case of standard weak convergence defined on the probability measures $\Delta(S)$.

With this notation, we are ready to define our key operator. Let $\Phi_t^{(0)}$ be the identity operator and for $\tau > 0$ define recursively the following:

$$\Phi_t^{(\tau)}(f,h) = \Phi_t(\Phi_{t+1}^{(\tau-1)}(f^2,h^2)). \tag{3}$$

Then, let define:

$$B_t^{(1)}(f,h) := (\Phi_t(f,h), f^2, h^2),$$

and then for $\tau > 1$, define recursively:

$$B_t^{(\tau)}(f,h) := B_t^{(1)} \circ B_{t+1}^{(\tau-1)}(f^2,h^2). \tag{4}$$

where $B_t^{(1)}$ is our shift operator that replaces the first element of the given sequence (f,g) with the appropriately chosen best-responses (values and policies), and similar notation then defines $B_t^{(\tau)}(f,h)$ with $1,2,\ldots,\tau$ elements replaced. Finally, define

$$\mathcal{B}_t = \bigcap_{\tau=1}^{\infty} B_t^{(\tau)}(\mathscr{V}).$$

This definition of $\mathcal{B}_t$ requires a few comments. Although our construction resembles the celebrated APS technique developed for repeated games (Abreu et al. (1986) and Abreu et al. (1990))), as well as a related approach in Mertens and Parthasarathy (1987) for stochastic games, it is different, however. First, we operate in *function spaces* rather then working with correspondences. For our approach, this difference is of utmost importance not only from that vantage point of how we prove existence, but notably also that our operator is designed specifically to capture our notion of time-consistency when studying (nonstationary) Markovian policies. Indeed, as argued by Doraszelski and Escobar (2012) and later by Balbus and Woźny (2016), such an specification of an self-generating set-iterative operator allows one to construct the set of *short-memory* (non-stationary) equilibria where the whole history is summarized by the state variable. Indeed, it is clear from the construction that current actions and future values *cannot* be conditioned upon *past* actions. This is a critical difference relative to standard APS-type constructions such as employed in quasi-hyperbolic discounting problems in Bernheim et al. (2015), for example. The same concerns relate to the selection of continuation values. Here, these are independent of both current (and past) states.

Second, in our construction, we use both sequences of values and *policies*. This is novel (but also see some recent contribution of Abreu et al. (2020a) and Abreu et al. (2020b)), and proves necessary in our problem due to the presence of time-consistency in dynastic preferences. Indeed, as the principle of optimality does not work in our case, it is not sufficient to characterize the future paths (and current generation preferences) using sequence of values only.

Third, as will be evident in the statement and proof of the main result, our construction involves *two* set-valued iterations. The first one is necessary to define incentive compatibility (here, to define best-responses). Indeed, as the preferences of each generation depend on the whole sequence of future values and actions, they cannot be summarized via the auxiliary game for some single continuation value. Moreover, since we seek for perfect equilibria but the game is infinite-horizon we cannot iterate from the last period. Here, we solve this problem but considering a sequence of composed best-responses to the give sequence of equilibrium candidates τ period ahead in (3). Then we take the limit with τ and obtain the infinite sequence of best-responses shifted till infinity. This allows us to consider $\mathcal{B}_t$, the composed best response of generation t, starting from some large set of candidate equilibrium objects. The second set-valued iteration considers sequence of $\{\mathcal{B}_t\}$, where each element is composed of *infinite* sequence of equilibrium candidates, and in particular, allows to show non-emptiness of $\mathcal{B}_1 \subset \mathscr{V}$, from which any selection is a MPE (policy and value) sequence.

We now state and discuss our assumptions.

Assumption 1 (Preferences)**.** *Assume for any t, V_t has the form:*

$$V_t(s, y_1, y_2, \ldots) := G_t(s, K_t(y_1, y_2, \ldots))$$

where $G_t : S \times [a, b] \mapsto \mathbb{R}$ for some real $a < b$, $s \in S$, $y_\tau \in [a, b]$ for any $\tau \in \mathbb{N}$, $K_t : [a, b]^\infty \mapsto [a, b]$ are both continuous functions satisfying the following conditions:

(i) G_t is increasing in both arguments satisfying $\max\limits_{s \in S, y \in [a,b]} |G_t(s, y)| \leq \gamma$ for some $\gamma > 0$;

(ii) for any $h > 0$, $y_1 > y_2$ the function $D_t^h : [0, \bar{s} - h] \mapsto \mathbb{R}$ defined as

$$D_t^h(s) := G_t(s, y_1) - G_t(s + h, y_2)$$

is a strictly single crossing function.[11]

Assumption 1 is rather standard excepting the fact that it allows for time-varying aggregators It implies the existence of a very general class of recursive and non-stationary preferences. The two conditions we impose, though, require a comment. We start from (ii). It requires that D_t^h is a strictly single crossing function for any h and continuation y. It implies strict single crossing property between savings i_t and the state s_t. It is satisfied , for example, whenever the current period date t utility is strictly concave in c_t , for example.[12] It is worth stressing, the strictly single crossing property is a ordinal condition in this recursive aggregator setting, and hence rather weak. The reason it is appropriate to our

[11] I.e., for any $s_2 > s_1$ we have $D_t^h(s_1) \geq 0$ implies $D_t^h(s_2) > 0$.

[12] Similar assumptions are imposed in Balbus et al. (2015a, 2020a) for a class of OLG models.

dynamic (and cardinal) problem results from the assumed aggregative structure of V_t, as given by K_t, as well as the way SSCP is imposed on the utility from the current consumption, and hence between i_t and s_t. These conditions guarantee in our case that in any MPE, each generation t investment policy $h_t : S \mapsto S$ is monotone increasing.

Assumption 2 (Transition)**.** *Assume for any t:*

(i) $q_t : S \mapsto \Delta(S)$ *is a measurable transition probability;*

(ii) $q_t(\cdot|s)$ *is a nonatomic probability measure for any* $s \in S$*;*

(iii) if $s_1 > s_2$ *then* $q_t(\cdot|s_1)$ *stochastically dominates* $q_t(\cdot|s_2)$*;*

(iv) q_t *has Feller property i.e.*

$$s \in S \mapsto \int_S \varphi(s')q_t(ds'|s)$$

is continuous whenever $\varphi : S \mapsto \mathbb{R}$ *is.*

Few comments are in order. Assumptions (i), *(ii)* and *(iii)* are standard. Point *(ii)* is necessary to assure our method is well defined. Typical examples of transitions satisfying 2 include models with multiplicative or additive shocks. To see that consider a production function f_t with $s_{t+1} = f_t(z_t, i_t)$ where z_t is a random shock with nonatomic distribution π_t, and for ensuring (ii) we assume f_t is injective in z. For a Borel set B, we can rewrite this transition process as:

$$q_t(B|i_t) = \int_S \mathbf{1}_B(f_t(z, i_t))\pi_t(dz).$$

It is now clear that our assumption are satisfied whenever, for example, $f_t(z_t, i_t) = z_t g_t(i_t)$ or $f_t(z_t, i_t) = g_t(i_t) + z_t$, for some continuous and increasing g_t, and some non-atomic probability measure π_t representing technology shocks. So this assumption in many applications is not very restrictive. It should be noted, though, that without assumption (ii), examples can be constructed where the argmax operator in (2) may not be well-defined.[13]

[13] See Example 1 in Balbus et al. (2016), for example.

Our final assumption imposes joint continuity of the certainty equivalence operator: that is,

Assumption 3 (Certainty Equivalent). *Suppose $\mu \in \Delta(S)$ is nonatomic. Let $f_n \Rightarrow f$ in $\mathscr{F}$, $\mu_n \Rightarrow \mu$ in $\Delta(S)$, and let $i_n \to i$ in S. Then*

$$\mathcal{M}_{i_n,t}(f_n, \mu_n) \to \mathcal{M}_{i,t}(f, \mu).$$

Notice the typical examples of certainty equivalence operators from the recursive utility literature satisfy these conditions (e.g., expected utility, CE given by integrals, quasi-linear means or entropic risk-measures, etc.)[14]

Now, let $\mathscr{E}$ be the set of all Markov perfect equilibria. Then, we have the main existence theorem of the paper:

Theorem 1. *Assume 1, 2 and 3. Then:*

(i) Any of $B_t^{(\tau)}(\mathscr{V})$ is weakly compact and

$$B_t^{(\tau)}(\mathscr{V}) \subset B_t^{(\tau-1)}(\mathscr{V}) \subset \ldots B_t^{(2)}(\mathscr{V}) \subset B_t^{(1)}(\mathscr{V}).$$

(ii) $\mathcal{B}_1$ is nonempty;

(iii) The following equality holds: $\mathscr{E} = \mathcal{B}_1$.

This is our main result. It proves *existence* of a Markov perfect equilibrium that are time-consistent solutions in the general class of altruistic consumption-savings/growth economies we study in this paper. Indeed, by definition for any selection from $\mathcal{B}_1$, say (f, h), we have $(f^2, h^2) \in \mathcal{B}_2$ and so on with $(f^t, h^t) \in \mathcal{B}_t$. This is the essence of time-consistency. As typically in self-generation arguments such equilibrium (selection) is non-stationary. In our case with time-variant preferences it is inevitable. All nonstationarity of equilibrium policies is, however, mapped into time index in our case. Indeed, Markov equilibria that we study do not depend on past actions or past draws of states. This is potentially restrictive per efficiency but justified under non-atomic noise condition. Second, MPE

[14] See also Kreps and Porteus (1978), Epstein and Zin (1989), Le Van and Vailakis (2005) or Bloise and Vailakis (2018), Balbus (2020) more recently.

exists with each policy h_t *in a class of increasing and (with loss of generality) upper-semicontinuous investment functions.* Such characterization regarding consumption policies is generally not available. Examples can be constructed with equilibrium consumption being non-monotone functions of states. Third, under our assumptions there are no MPE outside $\mathscr{V}$. Indeed, according to assumption 1 (ii) best-response map returns increasing investment policies for any continuation values (and policies). Fourth, our result is constructive in the sense it proposes an *iterative procedure* that converge to the *whole* set of MPE is our space of functions. Numerical algorithms can be proposed to compute this sequence of sets and approximate its limit. We now present the proof of theorem 1 with some technical lemmas moved to the appendix.

Proof. Identifying any element of $\mathscr{V}$ with a sequence of probability measures we conclude $\mathscr{V}$ is weakly compact. By Lemma 4 it follows that any of $B_t^{(\tau)}$ is a continuous operator on $\mathscr{V}$, hence $B_t^{(\tau)}(\mathscr{V})$ is weakly compact. Now we show the set $B_t^{(\tau)}(\mathscr{V})$ prefixed by $\mathscr{V}$ is descending (set inclusion order) in τ. Obviously, $B_t^{(1)}(\mathscr{V}) \subset \mathscr{V} = B_t^{(0)}$ for any τ and suppose it is true for some τ, that is

$$B_t^{(\tau)}(\mathscr{V}) \subset B_t^{(\tau-1)}(\mathscr{V}) \tag{5}$$

for any integer t. In particular we can substitute t by $t+1$ into (5). By equation (4) we easily conclude that (5) is true for $\tau + 1$ as well. Hence for any $B_t^{(\tau)}(\mathscr{V})$ is descending in τ. As a result, $\mathcal{B}_t$ is nonempty and weakly compact. We show the inclusion $\subset$. If $(f^*, h^*) \in \mathscr{V}$. Then by definition of $\mathcal{E}$ and B_1^τ

$$(f^*, h^*) \in B_1^{(\tau)}((f^*)^{(\tau)}, (h^*)^{(\tau)}) \in B_1^{(\tau)}(\mathscr{V})$$

for any τ, hence $(f^*, h^*) \in \mathcal{B}_1$. Now we show $\supset$. Let $(f^*, h^*) \in \mathcal{B}_1$. For any τ there exists then $(\tilde{f}_\tau, \tilde{h}_\tau) \in \mathscr{V}$ such that

$$(f^*, h^*) = B_1^{(\tau)}(\tilde{f}_\tau, \tilde{h}_\tau)$$

By (4) we have

$$(\tilde{f}_\tau, \tilde{h}_\tau) = B_1^{(1)}(B_2^{(\tau-1)}(\tilde{f}_\tau, \tilde{h}_\tau))$$

and by definition of the operator $B_1^{(1)}$ we have

$$(f_1^*, h_1^*) = T_1(B_2^{(\tau-1)}(\tilde{f}_\tau, \tilde{h}_\tau))$$

and

$$((f^*)^2, (h^*)^2) = B_2^{(\tau-1)}(\tilde{f}_\tau, \tilde{h}_\tau).$$

Furthermore:

$$\begin{aligned}
(f_2^*, h_2^*) &= T_2(B_3^{(\tau-2)}(\tilde{f}_\tau, \tilde{h}_\tau)), \\
((f^*)^3, (h^*)^3) &= B_3^{(\tau-2)}(\tilde{f}_\tau, \tilde{h}_\tau).
\end{aligned}$$

More generally for $k = 1, 2, \ldots, \tau$ we have:

$$\begin{aligned}
((f^*)^k, (h^*)^k) &= B_k^{(\tau-k+1)}(\tilde{f}_\tau, \tilde{h}_\tau). \\
(f_k^*, h_k^*) &= B_k(B_{k+1}^{(\tau-k)}(\tilde{f}_\tau, \tilde{h}_\tau)).
\end{aligned}$$

Hence for $k = 1, 2, \ldots, \tau$ we have

$$(f_k^*, h_k^*) = B_k((f^*)^{k+1}, (h^*)^{k+1}).$$

Since τ is arbitrary, we have the thesis. $\square$

In the above results we have proven existence of the nonstationary Markov perfect equilibrium. In certain cases, driven by numerical or theoretical considerations, the authors seek further characterization of the equilibrium strategies including its stationarity. For this reason, we now specify our model to address such questions.

4 Periodic model and periodic solutions

We now assume our model takes a special periodic form, and construct periodic Markov Perfect Equilibrium. We begin with the following specialization of assumptions 1, 2, and 3:

Assumption 4 (Periodic model)**.** *Assume there exists T such that for any integer t,* *(i)* $G_{t+T} = G_t$ *and* $K_{t+T} = K_t$*;*

(ii) $q_{t+T} = q_t$*;*

(iii) $\mathcal{M}_{t+T} = \mathcal{M}_t$

Under assumption 4, the dynastic altruistic preferences, the state transition structure, and the certainty equivalent will each have a periodic structure. Notice, additionally, one special case of assumption 4 is *time-invariance*, which is the case where $T = 1$.

We say that a Markov perfect equilibrium is *periodic* whenever $h_t = h_{t+T}$ for some T and any t. We now have our next key result in the paper.

Theorem 2. *Assume 1, 2, 3 and 4. Then, there exist a periodic Markov Perfect Equilibria. In particular, if the primitive data of the model is time-invariant* $(T = 1)$*, then there is a stationary Markov Perfect equilibrium.*

This result is important for a number of reasons. First, periodic equilibria are fixed points of B_1^T (i.e. fixed points of the t-th orbit of our operator). Their existence is hence a consequence of similar reasoning and lemmata as used in the proof of the main theorem in the previous section. Second, as recently shown by Berg (2017) and Berg and Kitti (2019), such periodic equilibria (or in their language, "elementary subpaths") are important from both theoretical and numerical perspective. We refer to reader to there papers for a discussion. Finally, we have as a special case of existence of *stationary* MPE for the time-invariant version of our model (but not necessarily stationary) version model. We believe that this is the most general existence result within the unified methodological framework that is known in the literature (but also see the related results in Balbus et al. (2020b)).

We finish this section with the proof of theorem 2 and discuss applications and relations to the literature in the final section.

Proof. For any finite sequence $(f_1, h_1, f_2, h_2 \ldots, f_T, h_T)$ define

$$J(f_1, h_1, f_2, h_2, \ldots, f_T, h_T) = (J_t(f_1, h_1, f_2, h_2 \ldots, f_T, h_T))_{t=1}^{\infty}$$

where $J_t(f_1, f_2, \ldots, f_T) = f_{t \bmod T} = f_{t-\lfloor \frac{t}{T} \rfloor T}$ and for $(f, h) \in \mathscr{V}$ we define

$$\Pi_T(f, h) = (f_1, h_1, f_2, h_2, \ldots, f_T, h_T).$$

Observe that $B_t^\tau = B_{t+T}^\tau$. Then the periodic distribution is $J(f_1^*, \ldots, f_T^*)$, where $(f_1^*, f_2^*, \ldots, f_T)$ is a fixed point of

$$(f_1, h_1, f_2, h_2 \ldots, f_T, h_T) \in (\mathscr{F} \times \mathscr{I})^T \mapsto \Pi_T(B_1^T J(f_1, \ldots, f_T)).$$

By Lemma 4 this operator is continuous. Obviously J and Π_T are continuous transformations too. Hence and by Schauder-Tychonoff Theorem, there exists a fixed point $(f_1^*, h_1^*, f_2^*, h_2^*, \ldots, f_T^*, h_T^*)$ of the operator above. □

5 Extensions toward paternalistic models

In this section, we show how our method can be extended towards models with direct paternalistic altruistic features. We do it by allowing the aggregator V_t to depend directly on the sequence of future generations' *consumption* policies h^t. For any $\mu \in \Delta(S)$ and Borel measurable and bounded $f : S \mapsto \mathbb{R}$, let $\hat{\mathcal{N}}$ be the Certainty Equivalent Operator used to evaluate the consumption policies of the future generations. To reduce computational burden, and for expositional simplicity, in this section we assume $\hat{\mathcal{N}}$ is time-invariant. But it should be clear from our argument, the generalization towards the case of time-varying $\hat{\mathcal{N}}$ is straightforward.

As before, let $\mathcal{N}_{i,t}(\cdot, h^{t+1,\tau-1}) = \hat{\mathcal{N}}\left(\cdot, q_t^{\tau+1}(\cdot|i, h^{t+1,\tau-1})\right)$, with $\mathcal{N}_{i,t}(\cdot) = \hat{\mathcal{N}}(\cdot, q_t(\cdot|i))$. We would need to adapt slightly our notation and assumptions from the previous section.

Assumption 5 (Preferences)**.** *Assume for any t, V_t has the form:*

$$V_t(s_1, s_2, y_2, s_3, y_3, \ldots) := G_t(s_1, K_t(s_2, y_2, s_3, y_3, \ldots))$$

where $G_t : S \times [a, b] \mapsto \mathbb{R}$ for some real $a < b$, and $K_t : (S \times [a, b])^\infty \mapsto [a, b]$ are both continuous functions satisfying the following conditions:

(i) *G_t is increasing in both arguments satisfying* $\max_{s\in S, y\in[a,b]} |G_t(s,y)| \leq \gamma$ *for some* $\gamma > 0$;

(ii) *for any* $h > 0$, $y_1 > y_2$ *the function* $D_t^h : [0, \bar{s} - h] \mapsto \mathbb{R}$ *defined as*

$$D_t^h(s) := G_t(s, y_1) - G_t(s + h, y_2)$$

is a strictly single crossing function.

We now adapt the definition on T and H to this new extended version of model as follows. Specifically, for $(f, h) \in (\mathscr{F} \times \mathscr{I})^\infty := \mathscr{V}$ let:

$$\hat{T}_t(f,h)(s) = \max_{i\in[0,s]} V_t \Big(s - i, \mathcal{N}^{\xi}_{i,t}(h_2), \mathcal{M}_{i,t}(f_2), \mathcal{N}^{\xi}_{i,t}(h_3, h^{2,0}), \mathcal{M}_{i,t}(f_3, h^{2,0}), ...$$

$$\ldots, \mathcal{N}^{\xi}_{i,t}(h_{\tau+1}, h^{2,\tau-2}), \mathcal{M}_{i,t}(f_{\tau+1}, h^{2,\tau-2}), \ldots\Big)$$

for each $s \in S$ and $t \in \mathbb{N}$, where $\xi := (\xi_t)_{t=1}^\infty$ is a fixed sequence of functions from $\mathscr{V}$, $h \in \mathscr{I}$ and $\mathcal{N}^{\xi}_{i,t}$ is used to evaluate future consumption policies, say $h_{t+\tau}$ and composed with $\xi_{t+\tau}$ to represent current direct preferences towards consumption τ period ahead. More formally, we have:

$$\mathcal{N}^{\xi}_{i,t}(h_{t+\tau}, h^{t+1,\tau-2}) := \mathcal{N}_{i,t}(\xi_{t+\tau} \circ h_{t+\tau}, h^{t+1,\tau-2}).$$

As before, to simplify notation, we assume an evaluation $\xi := (\xi_t)_{t=1}^\infty$ that is nonstationary but time-invariant. Similarly define:

$$\hat{H}_t(f,h)(s) = \arg\max_{i\in[0,s]} V_t \Big(s - i, \mathcal{N}^{\xi}_{i,t}(h_2), \mathcal{M}^{\xi}_{i,t}(f_2),$$

$$\mathcal{N}^{\xi}_{i,t}(h_3, h^{2,0}), \mathcal{M}_{i,t}(f_3, h^{2,0}), \ldots, \mathcal{N}^{\xi}_{i,t}(h_{\tau+1}, h^{2,\tau-2}), \mathcal{M}_{i,t}(f_{\tau+1}, h^{2,\tau-2}), \ldots\Big).$$

Theorem 3. *Suppose Assumptions 2 and 5 are satisfied. Moreover, Assumption 3 is satisfied for both $\mathcal{M}$ and $\mathcal{N}$. Then:*

(i) *Any of $B_t^{(\tau)}(\mathscr{V})$ is weakly compact and*

$$B_t^{(\tau)}(\mathscr{V}) \subset B_t^{(\tau-1)}(\mathscr{V}) \subset \ldots B_t^{(2)}(\mathscr{V}) \subset B_t^{(1)}(\mathscr{V}).$$

(ii) *$\mathcal{B}_1$ is nonempty;*

(iii) The following equality holds: $\mathscr{E} = \mathcal{B}_1$.

If additionally Assumption 4 holds, then there exist a periodic MPE. In particular, if the model is time-invariant ($T = 1$) *then there is a stationary equilibrium.*

Proof. Observe that the proof is almost the same as the proof of Theorem 1. We only need to adapt Lemma 4 to the assumptions of this theorem. Next we can continue the procedure from Theorem 1. To show the counterpart of Lemma 4 we show that if $(f_n, h_n) \to (f, g)$ in $(\mathscr{V} \times \mathscr{I})^\infty$ as $n \to \infty$, then

$$\hat{T}_t(f_n, h_n)(s) \to \hat{T}_t(f, h)(s) \text{ for any } s \in S \quad \text{and} \quad \hat{H}_t(f_n, h_n) \Rightarrow \hat{H}_t(f, h).$$

as $n \to \infty$. Let

$$\hat{\kappa}(i, f, h) := K_t \Big(\mathcal{N}^{\xi}_{i,t}(h_2), \mathcal{M}_{i,t}(f_2), \mathcal{N}^{\xi}_{i,t}(h_3), \mathcal{M}_{i,t}(f_3, h_2), \\ \ldots, \mathcal{N}^{\xi}_{i,t}(h_{\tau+1}, h^{2,\tau-2}), \mathcal{M}_{i,t}(f_{\tau+1}, h^{2,\tau-2}), \ldots \Big).$$

By definition of $\hat{\kappa}$ we have

$$\hat{T}_t(f, h)(s) = \max_{i \in [0,s]} G_t(s - i, \hat{\kappa}(i, f, h)).$$

Combining Assumptions 1, 3 (applied for both $\mathcal{M}$ and $\mathcal{N}$) and Lemma 1 we have the joint continuity of $\hat{\kappa}$. Hence and by Assumption 1 and Berge Maximum Theorem we have that $T'_t(f, h)(s)$ is continuous in all three arguments. Hence $\hat{T}_t(f_n, h_n)(s) \to T(f, h)(s)$ whenever $(f_n, h_n) \to (f, h)$ in $\mathscr{V}$. Moreover, $\hat{H}_t(f, h)(s)$ have closed graph i.e. if $(f_n, h_n) \Rightarrow (f, h)$ in $\mathscr{F}$, $s_n \to s$, $i_n \to i$ as $n \to \infty$ such that $i_n \in \hat{H}_t(f_n, h_n)(s_n)$ for all n then $i \in \hat{H}_t(f, h)(s)$. But from Lemma 3 we have $\hat{H}_t(f_n, h_n) \Rightarrow \hat{H}_t(f, h)$. □

6 Applications and related results

Example 1 (Separable non-paternalistic altruism and behavioral discounting)**.** Consider the special case of time separable aggregators:

$$V_t(c_t, U^t) = u(c_t) + \sum_{\tau=1}^{\infty} \beta_t^{\tau} U_{t+\tau},$$

where β_t^τ is a weight placed by generation t on the utility of the generation following τ periods ahead. Solving this model recursively, we obtain the following expression for the effective weight α_t^τ placed by generation t relative to the instantaneous utility τ periods ahead: it is given recursively by:

$$\alpha_t^\tau = \sum_{n=1}^{\tau} \beta_{t+\tau-n}^n \alpha_t^{\tau-n} \tag{6}$$

with initial $\alpha_t^0 = 1$. As a result, the preferences of generation t can be alternatively written as:

$$u(c_t) + \sum_{\tau=1}^{\infty} \alpha_t^\tau u(c_{t+\tau}). \tag{7}$$

The above expression allows one to see our model is general, and includes many behavioral discounting models in the existing literature as special cases. Specifically, for a given sequence of effective discount factors $\{\alpha_t^\tau\}$, we can compute the implied sequence of $\{\beta_t^\tau\}$ parameters. Solving the equation (6) recursively at each step, we then can determine $\beta_{t+\tau-n}^n$. Using this method, we can solve for the sequence $\{\beta_t^\tau\}$ for the quasi-hyperbolic or hyperbolic cases. Indeed, for example, for a time-invariant quasi-hyperbolic ($\beta-\delta$) discounting, we have: $\beta_t^\tau = \beta\delta^\tau(1-\beta)^{\tau-1}$. The results of our Theorems 1 and 2 can be therefore be applied. See also Galperti and Strulovici (2017) for a related derivation of quasi-hyperbolic discount factors,[15] Balbus et al. (2020b) for a recent reference on solving behavioral discounting models, and Saez-Marti and Weibull (2005) for further discussion of the relationship between discounting and altruism.

Example 2 (Endogenous intergenerational disagreement and parental transfers)**.** In the above examples, the weights α_t^τ placed on the value (or implied β_t^τ placed on utility) of the successor generations were fixed. The literature on dynamic behavioral models of choice often studies interesting cases where such weights are themselves endogenous, however. For example, in models with magnitude effects, such weights can be state and/or investment dependent. Models with investment dependent weights can be easily be seen to fit into our framework and assumptions. The case of state dependent weights (or discount factors) can also be

[15] Recall, there model satisfies altruism-stationary according to their axiom 8.

shown to fit into our framework but our key SSCP condition has to be strengthen to the cardinal case, namely strictly increasing differences (to assure that the best response is increasing for every continuation value).We now illustrate such environment in an altruistic OLG example, inspired by Pavoni and Yazici (2016), where the authors study the role of intergenerational disagreement on parental transfers. So let generation t preferences, endowed with state s_t and consuming c_t, by given by:

$$U_t = u(c_t) + \beta U_{t+1} + \delta(s_t)U_{t+2}.$$

That is, we assume each generation lives for one period, but derives utility from own consumptions and utility of the two consecutive offsprings (children and grandchildren). The weight placed on the children utility is β and the weight on grandchildren is given by $\delta(s_t)$ (and hence, is state dependent). We assume $\delta : S \to [0, 1)$ is monotone which implies that the richer the grandparents, the more they care about their grandchildren.

Although the current generation preferences are fixed (with β and $\delta(s_t)$), when making the investment decision each generation observes the weight placed on the following generation grandchildren is endogenous and under stochastic monotonicity assumptions placed on q, investment dependent. Indeed, when one generation (say, generation t) invests a lot, he/she makes its kids (generation $t + 1$) richer, but also makes the kids care more about their grandchildren (generation $t + 3$, for which generation t does not care directly). This creates a conflict or intergenerational disagreement on the optimal level of parental transfers. Moreover, observe the current example results in time-consistent intergenerational preferences only when $\delta(\cdot) = 0$. Hence, the higher the δ the higher the departure from time-consistent benchmark. This creates a friction between the direct benefit from investment and inheritance decisions of the following generations via time-inconsistency problem.

Under our stochastic setting the problem of each generation is then to choose $i_t \in [0, s_t]$ to maximize:

$$u(s_t - i_t) + \int_S \left[\beta f_{t+1}(s') + \delta(s_t) \int_S f_{t+2}(s'')q(ds''|h_{t+1}(s')) \right] q(ds'|i_t),$$

where, as in the general model, f_{t+1} (or f_{t+2}) is the considered utility of the following generation $t+1$ (or $t+2$ respectively) and h_{t+1} is the considered (investment) strategy of the following generation. It is straightforward to show that, whenever the period utility is strictly concave, and δ and q are increasing, then the objective has strict increasing differences between i_t and s_t, and our results (of theorems 1 and 2) hold.

We now present two examples involving paternalistic features and hence applications of our construction from section 5.

Example 3 (Consistent solutions in collective household models)**.** We now consider a dynamic maximization problem of a collective household with two individuals. Instantaneous utility functions of both individuals are given by continuous, strictly increasing and strictly concave u^1 and u^2 and we assume both discount the future utility streams exponentially with discount factors respectively given by δ_1 and δ_2. The first individual is more patient with $1 > \delta_1 > \delta_2 > 0$. The weights of both individuals utilities in the collective household preferences are $\eta_1 > 0$ and $\eta_2 > 0$. The utility of the collective household is then the following:

$$E_s \left\{ \sum_{t=0}^{\infty} \eta_1 \delta_1^t u^1(c_{1,t}) + \eta_2 \sum_{t=0}^{\infty} \delta_2^t u^2(c_{2,t}) \right\},$$

where $c_{i,t}$ is the i-th individual consumption in period t. Clearly, $c_{1,t}+c_{2,t} \leq s_t$, and we assume that s_t is a Markov chain controlled by $(c_{1,t}, c_{2,t})_{t=0}^{\infty}$, such that state s_{t+1} is drawn from distribution $q(\cdot|s_t - c_{1,t} - c_{2,t})$. See Gollier and Zeckhauser (2005), Zuber (2011) or Jackson and Yariv (2015) for motivation of studying such problems.

We now focus on a dynastic representation of this collective household problem with each (time-invariant) generation t preferences given by:

$$E_s \sum_{\tau=0}^{\infty} (\delta_1)^{\tau} \left(\eta_1 u^1(c_{1,t+\tau}) + \eta_2 \left(\frac{\delta_2}{\delta_1} \right)^{\tau} u^2(c_{2,t+\tau}) \right).$$

We will show how the tools from our paper can be used to characterize MPE in this problem also. In the collective household problem, MPE is given by the sequence $(h_{1,t}, h_{2,t}, h_t)_{t=0}^{\infty}$ of measurable functions $h_{i,t} : S \to S$ and $h_t : S \to S$

such that for any t we have: $h_{1,t}(s) + h_{2,t}(s) + h_t(s) = s$. We start with an important observation:

Proposition 1. *Suppose $(h_{1,t}, h_{2,t}, h_t)_{t=0}^{\infty}$ is a MPE. There exist measurable functions $\beta : S \to S$ and $\gamma : S \to S$ such that:*

- *for each t and each $s \in S$: $h_{1,t}(s) = \beta(c_t(s))$ and $h_{2,t}(s) = \gamma(c_t(s))$ where $c_t(s) = h_{1,t}(s) + h_{2,t}(s) = s - h_t(s)$.*
- *moreover β and γ are such that for each $c \in S$:*

$$\eta_1 u^1(\beta(c)) + \eta_2 u^2(\gamma(c)) = \max_{c_1, c_2 \geq 0} \{\eta_1 u^1(c_1) + \eta_2 u^2(c_2)\} \text{ s.t. } c_1 + c_2 \leq c.$$

Proof. Let $(h_{1,t}, h_{2,t}, h_t)_{t=0}^{\infty}$ be a MPE. Evaluate

$$w(s) := \sum_{t=1}^{\infty} (\delta_1)^t E_s \left(\eta_1 u^1(h_{1,t}(s_t)) + \eta_2 \left(\frac{\delta_2}{\delta_1}\right)^t u^2(h_{2,t}(s_t)) \right),$$

where E_s is take with resect to realisation of s_{t+1} governed by $Q(\cdot|h_t(s_t))$. Now suppose by contradiction that for some t consumptions $h_{1,t}, h_{2,t}$ are not solving

$$\max_{c_1, c_2 \geq 0} \{\eta_1 u^1(c_1) + \eta_2 u^2(c_2)\} \text{ s.t. } c_1 + c_2 \leq s - h_t(s),$$

for some s. That is, there exists s and c_1^*, c_2^* such that $c_1^* + c_2^* \leq s - h_t(s)$ and

$$\eta_1 u^1(c_1^*) + \eta_2 u^2(c_2^*) > \eta_1 u^1(h_{1,t}(s)) + \eta_2 u^2(h_{2,t}(s))$$

This implies that:

$$\begin{aligned} &\eta_1 u^1(c_1^*) + \eta_2 u^2(c_2^*) + \delta_1 \int_S w(s') Q(ds'|s - h_t(s)) > \\ &\eta_1 u^1(h_{1,1}(s)) + \eta_2 u^2(h_{2,1}(s)) + \delta_1 \int_S w(s') Q(ds'|s - h_t(s)). \end{aligned}$$

Since c_1^* and c_2^* are feasible, this means that $(h_{1,t}, h_{2,t}, h_t)_{t=0}^{\infty}$ cannot be a MPE. □

The above lemma allows us to rewrite consumptions in any MPE $(h_{1,t}, h_{2,t}, h_t)_{t=0}^{\infty}$ in the collective household problem using the following substitution: $h_{1,t}(s) = \beta(s - h_t(s))$ and $h_{2,t}(s) = \gamma(s - h_t(s))$, where

$$\eta_1 u^1(\beta(c)) + \eta_2 u^2(\gamma(c)) = \max_{c_1, c_2 \geq 0} \{\eta_1 u^1(c_1) + \eta_2 u^2(c_2)\} \text{ s.t. } c_1 + c_2 \leq c.$$

Next let us introduce the aggregate utility from the aggregate consumption:

$$u_t(c) := \eta_1 u^1(\beta(c)) + \eta_2 \left(\frac{\delta_2}{\delta_1}\right)^t u^2(\gamma(c)).$$

Then the problem of finding MPE in the collective household model can be hence reduced to finding investment $(h_t)_t$ that is a MPE of the dynastic game with preferences:

$$E_s \sum_{\tau=0}^{\infty} \delta_1^{\tau} u_{\tau}(s_{t+\tau} - i_{t+\tau})$$

and transition $q(\cdot|i_{t+\tau})$. In the view of proposition 3 this is a special case of our model (with non-stationary paternalistic utility $\xi_\tau := u_\tau$) and our result (proposition 3) hold. Related results under stronger conditions were obtained by Drugeon and Wigniolle (2016) (algebraic examples for Cobb-Douglas technology and CIES preferences) and Balbus et al. (2021) (obtained GEE under stochastic convexity of the transition q).

Example 4 (OLG with two-sided altruism and future bias)**.** Gonzalez et al. (2018) consider an OLG economy with altruism where each generation preferences are given by:

$$U_t = u^y(c_t^y) + u^o(c_{t+1}^o) + \mu U_{t-1} + \lambda U_{t+1}.$$

Here $\mu > 0$ is the weight placed to backward and $\lambda > 0$ towards forward altruism. Solving the system of equalities (with $\mu + \lambda < 1$) they obtain well-being of generation t as:

$$\sum_{\tau=1}^{\infty} \theta^{\tau}(u^y(c_{t-\tau}^y) + u^o(c_{t-\tau+1}^o)) + u^y(c_t^y) + u^o(c_{t+1}^o) + \sum_{\tau=1}^{\infty} \delta^s(u^y(c_{t+\tau}^y) + u^o(c_{t+\tau+1}^o)),$$

where θ (resp. δ) is the effective backward (resp. forward) discounting factor; both obtained from solving the system of equalities above[16]. Ignoring the part of preferences generation t cannot control we obtain

$$u^y(c_t^y) + \theta u^0(c_t^0) + \sum_{\tau=1}^{\infty} \delta^{\tau}[u^y(c_{t+\tau}^y) + \delta^{-1} u^o(c_{t+\tau}^o)].$$

[16] Here $\theta = \frac{1-\sqrt{1-4\mu\lambda}}{2\lambda}$ and $\delta = \frac{1-\sqrt{1-4\mu\lambda}}{2\mu}$. See page 440 in Gonzalez et al. (2018) but also Hori and Kanaya (1989).

Gonzalez et al. (2018) show for $\theta < 1 < \delta^{-1}$, these preferences exhibit forward bias (a form of quasi-hyperbolic discounting with present-bias > 1). They consider a sequence of short lived governments each aiming to maximize generation t preferences and seek for MPE of such intergenerational game. They consider CIES preferences, linear technology and focus on stationary MPE in linear strategies.

We now show how to map this problem into our model. Assume u^i is continuous, increasing and strictly concave. Recall, feasibility requires for any t: $c_t^0 + c_t^y + i_t \leq s_t$. Denote

$$u(c) := u^y(\beta(c)) + \theta u^0(\gamma(c)) = \max_{c^y, c^0 \geq 0} \{u^y(c^y) + \theta u^0(c^0)\} \text{ s.t. } c^y + c^0 \leq c,$$

where β and γ are the shares of the aggregate consumption c dedicated to young and old respectively. Similarly to lemma 1 above we can argue that in any MPE $h_t^y(s) = \beta(s - h_t(s))$ and $h_t^y(s) = \gamma(s - h_t(s))$. This allows us to write aggregate utility in period $t \geq 1$ as:

$$\tilde{u}(c) := u^y(\beta(c)) + \delta^{-1} u^0(\gamma(c)).$$

Hence finding MPE reduces to finding investment $(h_t)_t$ that is a MPE of the game with preferences given by:

$$u(s_t - i_t) + E_{s_t} \sum_{\tau=1}^{\infty} \delta^\tau \tilde{u}(s_{t+\tau} - i_{t+\tau}).$$

Again, this is a special case of our model with non-stationary paternalistic utility $\xi_t := \tilde{u}$) and our result (proposition 3) hold.

We finish this section by discussing a few examples (and a counterexample) illustrating the role of our assumptions placed on the certainty equivalent operator.

Example 5 (Endogenous disappointment aversion)**.** To exemplify the class of certainty equivalents, we consider a model that allows for endogenous preference formation in the form of (endogenous) disappointment aversion. Following Gul (1991), we construct the certainty equivalent $\hat{\mathcal{M}}$ as a unique solution of the equation:

$$\mathbb{R} \in \zeta = \int_S f(s')\mu(ds') - \delta \int_{s':f(s')<\zeta} (\zeta - f(s'))\mu(ds').$$

Here we show this certainty equivalent satisfies our continuity assumption.

Proposition 2. *If $f_n \Rightarrow f$, $\mu_n \Rightarrow \mu$ and μ is nonatomic, then $\mathcal{M}(f_n, \mu_n) \to \mathcal{M}(f, \mu)$.*

Proof. We show,

$$\Phi(\zeta, f, \mu) = \int_S f(s')\mu(ds') - \delta \int_{\{s': f(s') < \zeta\}} (\zeta - f(s'))\mu(ds')$$

is jointly continuous. We only need to show the continuity of

$$(\zeta, f, \mu) \mapsto \int_{\{s': f(s') < \zeta\}} (\zeta - f(s'))\mu(ds').$$

Let S_0 be the set of continuity points of f. Let $s' \in S$. If $f_n(s') < \zeta$ for all n, then $f(s') \leq \zeta$. Then

$$\mathbf{1}_{[0,\zeta_n)}(f_n(s'))(\zeta - f_n(s')) = f_n(s') - \zeta_n \to f(s') - \zeta.$$

If $f(s') < \zeta$ then $f(s') - \zeta = \mathbf{1}_{\{s' \in S: f(s') < \zeta\}}(f(s') - \zeta)$. If $f(s') = \zeta$ then

$$0 = f(s') - \zeta = (f(s') - \zeta)\mathbf{1}_{\{s' \in S: f(s') < \zeta\}}.$$

In both cases

$$\mathbf{1}_{[0,\zeta_n)}(f_n(s'))(\zeta - f_n(s')) \to \mathbf{1}_{[0,\zeta)}(f(s'))(\zeta - f(s'))$$

as $n \to \infty$. If $f_n(s') \geq \zeta$ for all n then $f(s') \geq \zeta$ as well. In such a case both side above are zeros at the same time. Applying Skorohod Representation Theorem again we conclude

$$\Phi(\zeta_n, f_n, \mu_n) \to \Phi(\zeta, f, \mu).$$

We show, the thesis. Let $\zeta_n = \hat{\mathcal{M}}(f_n, \mu_n)$ and $\zeta = \hat{\mathcal{M}}(f, \mu)$. By definition

$$\zeta_n = \Phi(\zeta_n, f_n, \mu_n).$$

The sequence ζ_n is commonly bounded since

$$\zeta_n = \Phi(\zeta_n, f_n, \mu_n) \leq \int_S f_n(s')\mu_n(ds') \to \int_S f(s')\mu(ds').$$

Hence we may suppose without loss of generality $\zeta_n \to \zeta$. Then

$$\zeta_n = \Phi(\zeta_n, f_n, \mu_n) \to \Phi(\zeta, f, \mu)$$

as $n \to \infty$ since Φ is jointly continuous. Since $\zeta_n \to \zeta$ a the same time, $\zeta = \Phi(\zeta, f, \mu)$. Since Φ is decreasing in the first argument, $\zeta = \hat{\mathcal{M}}(f, \mu)$ $\square$

Example 6 (Independence, Allais paradox and endogenous certainty equivalents)**.** We can also construct the certainty equivalent in the spirit of Dekel (1986). The lottery is identified with μf^{-1}. Following Dekel we find a fixed point of F

$$F(x) = \int_S w(f(s'), x)\mu(ds')$$

for some w- jointly continuous functions, increasing with the first argument and decreasing with respect to the second[17] such that $w(x, x) = x$ for $x \in \mathbb{R}_+$. That is $\hat{\mathcal{M}}(f, \mu)$ solves

$$\hat{\mathcal{M}}(f, \mu) = \int_S w\left(f(s'), \hat{\mathcal{M}}(f, \mu)\right)\mu(ds').$$

Similarly as in case of example 5, we conclude such equivalent obeys our assumptions.

Proposition 3. *$\hat{\mathcal{M}}(f, \mu)$ is increasing in both arguments whenever we restrict attention to increasing f. Assume μ is nonatomic. Then, if $\mu_n \Rightarrow \mu$ and $f_n \Rightarrow f$ then $\hat{\mathcal{M}}(f_n, \mu_n) \to \tilde{\hat{M}}(f, \mu)$.*

Proof. Let $\mu_1 \leq_{st} \mu_2$ and $f_1 \leq f_2$ pointwise and assume f_1, f_2 are both increasing. Then

$$\int_S w(f_1(s'), x)\mu_1(ds') \leq \int_S w(f_2(s'), x)\mu_2(ds').$$

Then

$$\begin{aligned} \hat{\mathcal{M}}(f_1, \mu_1) &= \int_S w(f_1(s'), \hat{\mathcal{M}}(f_1, \mu_1))\mu_1(ds') \\ \leq \int_S w(f_2(s'), \hat{\mathcal{M}}(f_1, \mu_1))\mu_2(ds') &= F\left(\hat{\mathcal{M}}(f_1, \mu_1)\right). \end{aligned} \tag{8}$$

Now suppose

$$\hat{\mathcal{M}}(f_2, \mu_2) \leq \hat{\mathcal{M}}(f_1, \mu_1).$$

Then by (8) we have

$$\hat{\mathcal{M}}(f_2, \mu_2) = F\left(\hat{\mathcal{M}}(f_2, \mu_2)\right) \geq F\left(\hat{\mathcal{M}}(f_1, \mu_1)\right) \geq \hat{\mathcal{M}}(f_1, \mu_1).$$

[17] Dekel requires the existence of a unique fixed point. It is obviously obtained if F is decreasing.

Hence we have $\hat{\mathcal{M}}(f_2, \mu_2) = \hat{\mathcal{M}}(f_1, \mu_1)$. Consequently we have the monotonicity restricted to increasing functions f. Now assume μ is nonatomic. Then repeating the same argument in case of example 5, we have the thesis. $\square$

Example 7 (Counterexample to quantile models)**.** Notably, there are certainty equivalents will do not satisfy our continuity assumption. One important counterexample concerns the certainty equivalents given by quantiles. So let $\zeta > 0$ be given and let

$$\hat{\mathcal{M}}(f, \mu) := \inf \{x > 0 : \mu \{s \in S : f(s) < x\} \leq \zeta \leq \mu \{s \in S : f(s) \leq x\}\}.$$

Let $f_n \Rightarrow f$ in $\mathscr{F}$ and $\mu_n \Rightarrow \mu$ in $\Delta(S)$. Let $x_n := \hat{\mathcal{M}}(f_n, \mu_n)$ and $x = \hat{\mathcal{M}}(f, \mu)$. We can then give a counterexample to $x_n \to x$.

Let $\zeta = 1/2$, $S = [0, 1]$ and let μ_n be a measure whose density is $\rho_n(s) = \left(1 - \frac{1}{n}\right) s^{-\frac{1}{n}}$. The distribution function is $F_n(s) = s^{1-\frac{1}{n}}$ for $s \in [0, 1]$. Clearly $F_n(s) \to s$ for $s \in [0, 1]$ hence $\mu_n \Rightarrow \mu$ where μ is a standard Lebesgue measure. Let

$$f_n(s) = \begin{cases} 2^n s^n & \text{for} \quad s \in [0, 1/2] \\ 1 & \text{for} \quad s \in [1/2, 1]. \end{cases}$$

Hence, the limit is $f_n \Rightarrow f$ where $f(s) = \mathbf{1}_{[1/2,1]}(s)$ for $s \in [0, 1]$. We find x such that

$$\mu \{s \in S : f(s) < x\} \leq \frac{1}{2} \leq \mu \{s \in S : f(s) \leq x\}.$$

We have for any $s' \in (0, 1)$

$$\mu \{s \in S : f(s) < s'\} = \mu \{s \in S : f(s) = 0\} = \frac{1}{2},$$

and

$$\mu \{s \in S : f(s) \leq s'\} = \mu \{s \in S : f(s) = 0\} = \frac{1}{2}.$$

Hence $x = \hat{\mathcal{M}}(f, \mu) = 0$. Furthermore, for $s' \in [0, 1/2)$

$$\{s \in S : f_n(s) < s'\} = \left\{s \in S : s < \frac{\sqrt[n]{s'}}{2}\right\} = \left[0, \frac{\sqrt[n]{s'}}{2}\right).$$

Hence

$$\mu_n \{s \in S : f_n(s) < s'\} = \mu_n \{s \in S : f_n(s) \leq s'\} = \mu_n \left(\left[0, \frac{\sqrt[n]{s'}}{2}\right)\right) = \left(\frac{\sqrt[n]{s'}}{2}\right)^{1-\frac{1}{n}}.$$

We find $s' \in (0, 1/2]$ such that

$$\left(\frac{\sqrt[n]{s'}}{2}\right)^{1-\frac{1}{n}} = \frac{1}{2} \Leftrightarrow s' = \frac{1}{2^{\frac{n}{n-1}}}.$$

Hence

$$x_n = \hat{\mathcal{M}}(f_n, \mu_n) = \frac{1}{2^{\frac{n}{n-1}}} \to \frac{1}{2} \neq 0 = x = \hat{\mathcal{M}}(f, \mu).$$

Interestingly, whenever the limiting measure is non-atomic the correspondence of certainty equivalents is upper-semicontinuous (has a closed graph). It is a continuous function only, when the limiting f is strictly increasing.

A Appendix

Lemma 1. *Let $h_n \Rightarrow h$ as $n \to \infty$ in $\mathcal{I}^\infty$. Then for any $i \in S$*

$$q_t^{\tau+1}(\cdot|i, h_n^{t+1,\tau-1}) \Rightarrow q_t^{\tau+1}(\cdot|i, h^{t+1,\tau-1}) \text{ as } n \to \infty. \tag{9}$$

Proof. We prove (9) by induction with respect to τ. If $\tau = 1$ then by Assumption 2 the thesis is done. Now suppose (9) holds for some $\tau - 1$ with $\tau > 1$ and any $t \in \mathbb{R}$. We show (9) holds for τ and any $t \in \mathbb{R}$. Let $\phi : S \mapsto \mathbb{R}$ be a continuous function and $i \in S$. Let

$$w(i) := \int_S \varphi(s') q_{t+1}(ds'|i), \tag{10}$$

and

$$\mu_n := q_{t+1}^{\tau}(\cdot|i, h_n^{t+2,\tau-1}) \text{ and } \mu := q_{t+1}^{\tau}(\cdot|i, h^{t+2,\tau-1}). \tag{11}$$

By induction hypothesis w is continuous and $\mu_n \Rightarrow \mu$. Moreover, μ is nonatomic, hence concentrated on the set of continuity points of $h_{t+\tau}$. Consequently there exist a probability space $(\Omega, \mathscr{B}, P)$ and a sequence of real valued random variables Y_n on Ω whose distribution is μ_n, a real valued random variable Y whose distribution is μ such that $Y_n(\omega) \to Y(\omega)$ for all $\omega \in \Omega$. Since μ is concentrated on the continuity points of $h_{t+\tau}$ hence

$$w(h_{t+\tau,n}(Y_n(\omega))) \to w(h_{t+\tau}(Y(\omega))) \text{ as } n \to \infty, \text{ for } P - a.a. \omega \in \Omega. \tag{12}$$

Combining (10), (11), (12) we have

$$\lim_{n\to\infty}\int_S w(h_{t+\tau,n}(s'))\mu_n(ds') = \lim_{n\to\infty}\int_\Omega w(h_{t+\tau,n}(Y_n(\omega)))P(d\omega)$$
$$= \int_\Omega w(h_{t+\tau}(Y(\omega)))P(d\omega) = \int_S w(h_{t+\tau}(s))\mu(ds)$$

and hence (9). Hence the proof is complete. □

Lemma 2. *The operator Φ maps $\mathscr{V}$ into itself.*

Proof. Let $(f,h) \in \mathscr{V}$ and $t \in \mathbb{N}$. By Assumption 1 we immediately have that $T_t(f,h)$ is increasing. We show that any element of H_t is increasing. For $i \in S$ put

$$\kappa(i) := K_t\left(\mathcal{M}_{i,t}(f_2), \mathcal{M}_{i,t}(f_3,h_2), \ldots, \mathcal{M}_{i,t}(f_\tau, h^{t+1,\tau}), \ldots\right).$$

By Assumption 1 $H_t(f,h)(s) = \arg\max_{i\in[0,s]} G_t(s-i,\kappa(i))$. By Assumption 1 and Proposition 1 in Balbus et al. (2015a) it follows that any selection of $H_t(f,h)(s)$ is increasing. □

Lemma 3. *Let Γ is a non-empty valued correspondence from $[\alpha,\beta]$ into a bounded interval in $\mathbb{R}$. Suppose Γ has closed graph and any selection is increasing function. Then:*

(i) if $\Gamma(x)$ is a singleton then any selection of Γ is continuous at x;

(ii) if $\Gamma(x)$ is not singleton, any selection of Γ is discontinuous at x.

Proof. Suppose $\Gamma(x) = \{y\}$ and let γ be any selection of Γ. Then if $x_n \to x$ as $n \to \infty$, then all cluster points of the sequence $\gamma(x_n)$ is in $\Gamma(x)$. But $\Gamma(x) = \{y\}$ is a singleton, hence $\gamma(x_n) \to y$ as $n \to \infty$. But $\gamma(x) \in \Gamma(x)$ as well, hence $\gamma(x_n) \to \gamma(x)$ as $n \to \infty$. Consequently γ is continuous. Now suppose $\Gamma(x)$ is not singleton. Let $y_1 < y_2$ and both belong to $\Gamma(x)$. Then any selection γ of Γ satisfies

$$\lim_{x'\uparrow x}\gamma(x') \le y_1 < y_2 \le \lim_{x'\downarrow x}\gamma(x'),$$

hence γ is discontinuous at x. □

Lemma 4. *If $(f_n, h_n) \to (f, h)$ in $\mathscr{V}$ then $T_t(f_n, h_n) \rightrightarrows T_t(f, h)$. Moreover, $\chi_n \rightrightarrows \chi$ where χ_n is a selection of $H_t(f_n, h_n)$ and χ is a selection of $H_t(f, h)$.*

Proof. Let us modify definitions from Lemma 2. Let

$$\kappa(i, f, h) := K_t\left(\mathcal{M}_{i,t}(f_2), \mathcal{M}_{i,t}(f_3, h_2), \ldots, \mathcal{M}_{i,t}(f_{\tau+1}, h^{2,\tau-2}), \ldots\right).$$

By definition of κ we have

$$T_t(f, h)(s) = \max_{i \in [0,s]} G_t(s - i, \kappa(i, f, h)).$$

Combining Assumptions 1, 3 and Lemma 1 we have the joint continuity of κ. Hence and by Assumption 1 and Berge Maximum Theorem we have that $T_t(f, h)(s)$ is continuous in all three arguments. Hence $T_t(f_n, h_n)(s) \to T(f, h)(s)$ whenever $(f_n, h_n) \to (f, h)$ in $\mathscr{V}$ and the more $T_t(f_n, h_n)(s) \rightrightarrows T(f, h)(s)$. Moreover, $H_t(f, h)(s)$ have closed graph i.e. if $(f_n, h_n) \rightrightarrows (f, h)$ in $\mathscr{F}$, $s_n \to s$, $i_n \to i$ as $n \to \infty$ such that $i_n \in H_t(f_n, h_n)(s_n)$ for all n then $i \in H_t(f, h)(s)$. But from Lemma 3 we have $H_t(f_n, h_n) \rightrightarrows H_t(f, h)$. □

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2019): "Education policy and intergenerational transfers in equilibrium," *Journal of Political Economy*, 127, 2569–2624.

Abel, A. B. (1987): "Operative gift and bequest motives," *The American Economic Review*, 77, 1037–1047.

Abreu, D., B. Brooks, and Y. Sannikov (2020a): "Algorithms for stochastic games with perfect monitoring," *Econometrica*, 88, 1661–1695.

——— (2020b): "A 'pencil-sharpening' algorithm for two player stochastic games with perfect monitoring," MS.

Abreu, D., D. Pearce, and E. Stacchetti (1986): "Optimal cartel equilibria with imperfect monitoring," *Journal of Economic Theory*, 39, 251–269.

——— (1990): "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica*, 58, 1041–1063.

Adams, A., L. Cherchye, B. De Rock, and E. Verriest (2014): "Consume now or later? Time inconsistency, collective choice, and revealed preference," *American Economic Review*, 104, 4147–83.

Akerlof, G. A. (1991): "Procrastination and obedience," *The American Economic Review*, 81, 1–19.

Ameriks, J., J. Briggs, A. Caplin, M. D. Shapiro, and C. Tonetti (2017): "The long-term-care insurance puzzle: modeling and measurement," *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, 110, 1–59.

——— (2020): "Long-term-care utility and late-in-life saving," *Journal of Political Economy*, 128, 2375–2451.

Asheim, G. B. (2010): "Global welfare comparisons," *Canadian Journal of Economics/Revue canadienne d'economique*, 43, 1412–1432.

Balbus, Ł. (2020): "On recursive utilities with non-affine aggregator and conditional certainty equivalent," *Economic Theory*, 70, 551–577.

Balbus, Ł., J.-P. Drugeon, and Ł. Woźny (2021): "On collective intertemporal choice, time-consistent decision rules and altruism," MS.

Balbus, Ł., A. Jaśkiewicz, A. S. Nowak, , and Ł. Woźny (2017): "A note on Markov perfect equilibria in a class of non-stationary stochastic bequest games," *Journal of Mathematical Analysis and Applications*, 456, 394 – 401.

Balbus, Ł., A. Jaśkiewicz, and A. S. Nowak (2015a): "Bequest games with unbounded utility functions," *Journal of Mathematical Analysis and Applications*, 427, 515 – 524.

——— (2016): "Non-paternalistic intergenerational altruism revisited," *Journal of Mathematical Economics*, 63, 27–33.

——— (2020a): "Equilibria in altruistic economic growth models," *Dynamic Games and Applications*, 10, 1–18.

Balbus, Ł., K. Reffett, and Ł. Woźny (2015b): "Time consistent Markov policies in dynamic economies with quasi-hyperbolic consumers," *International Journal of Game Theory*, 44, 83–112.

——— (2018): "On uniqueness of time-consistent Markov policies for quasi-hyperbolic consumers under uncertainty," *Journal of Economic Theory*, 176, 293 – 310.

——— (2020b): "Time consistent equilibria in dynamic models with recursive payoffs and behavioral discounting," MS.

Balbus, Ł. and Ł. Woźny (2016): "A strategic dynamic programming method for studying short-memory equilibria of stochastic games with uncountable number of states," *Dynamic Games and Applications*, 6, 187–208.

Barczyk, D. and M. Kredler (2020): "Blast from the past: the altruism model is richer than you think," MS.

Barro, R. J. (1974): "Are government bonds net wealth?" *Journal of Political Economy*, 82, 1095–1117.

Becker, G. S. and N. Tomes (1979): "An equilibrium theory of the distribution of income and intergenerational mobility," *Journal of Political Economy*, 87, 1153–1189.

Berg, K. (2017): "Extremal pure strategies and monotonicity in repeated games," *Computational Economics*, 49, 387–404.

Berg, K. and M. Kitti (2019): "Equilibrium paths in discounted supergames," *Discrete Applied Mathematics*, 260, 1 – 27.

Bernheim, B. D. and K. Bagwell (1988): "Is everything neutral?" *Journal of*

Political Economy, 96, 308–338.

Bernheim, B. D. and D. Ray (1986): "On the existence of Markov-consistent plans under production uncertainty," *Review of Economic Studies*, 53, 877–882.

Bernheim, B. D., D. Ray, and S. Yeltekin (1999): "Self-control, savings, and the low-asset trap," Stanford University.

——— (2015): "Poverty and self-control," *Econometrica*, 83, 1877–1911.

Bernheim, B. D., A. Shleifer, and L. H. Summers (1985): "The strategic bequest motive," *Journal of Political Economy*, 93, 1045–1076.

Bloise, G. and Y. Vailakis (2018): "Convex dynamic programming with (bounded) recursive utility," *Journal of Economic Theory*, 173, 118 – 141.

Boar, C. (2021): "Dynastic precautionary savings," *The Review of Economic Studies*, forthcoming.

Braun, R. A., K. A. Kopecky, and T. Koreshkova (2017): "Old, sick, alone, and poor: a welfare analysis of old-age social insurance programmes," *The Review of Economic Studies*, 84, 580–612.

Cao, D. and I. Werning (2018): "Saving and dissaving with hyperbolic discounting," *Econometrica*, 86, 805–857.

Chade, H., P. Prokopovych, and L. Smith (2008): "Repeated games with present-biased preferences," *Journal of Economic Theory*, 139, 157–175.

Chakraborty, A., Y. Halevy, and K. Saito (2020): "The relation between behavior under risk and over time," *American Economic Review: Insights*, 2, 1–16.

Chambers, C., F. Echenique, and A. A. Miller (2021): "Decreasing patience," MS.

Dasgupta, P. and E. Maskin (2005): "Uncertainty and hyperbolic discounting," *American Economic Review*, 95, 1290–1299.

De Nardi, M. (2004): "Wealth inequality and intergenerational links," *The Review of Economic Studies*, 71, 743–768.

De Nardi, M., E. French, and J. B. Jones (2016): "Savings after retirement: a survey," *Annual Review of Economics*, 8, 177–204.

DeJarnette, P., D. Dillenberger, D. Gottlieb, and P. Ortoleva (2020): "Time lotteries and stochastic impatience," *Econometrica*, 88, 619–656.

Dekel, E. (1986): "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom," *Journal of Economic Theory*, 40, 304–318.

Doepke, M. and G. Sorrenti (2019): *Love, money, and parenting: how economics explains the way we raise our kids*, Princeton, NJ: Princeton Univ. Press.

Doepke, M., G. Sorrenti, and F. Zilibotti (2019): "The economics of parenting," *Annual Review of Economics*, 11, 55–84.

Doepke, M. and F. Zilibotti (2017): "Parenting with style: altruism and paternalism in intergenerational preference transmission," *Econometrica*, 85, 1331–1371.

Doraszelski, U. and J. F. Escobar (2012): "Restricted feedback in long term relationships," *Journal of Economic Theory*, 147, 142–161.

Drugeon, J. P. and B. Wigniolle (2016): "On time consistent policy rules for heterogeneous discounting programs," *Journal of Mathematical Economics*, 63, 174–187.

Dziewulski, P. (2018): "Revealed time preference," *Games and Economic Behavior*, 112, 67 – 77.

Echenique, F., T. Imai, and K. Saito (2020): "Testable implications of models of intertemporal choice: exponential discounting and its generalizations," *American Economic Journal: Microeconomics*, 12, 114–43.

Epstein, L. G. and S. E. Zin (1989): "Substitution, risk aversion, and the remporal behavior of consumption and asset returns: a theoretical framework," *Econometrica*, 57, 937–969.

Farhi, E. and I. Werning (2007): "Inequality and social discounting," *Journal of Political Economy*, 115, 365–402.

Fels, S. and R. Zeckhauser (2008): "Perfect and total altruism across the generations," *Journal of Risk and Uncertainty*, 37, 187–197.

Feng, T. and S. Ke (2018): "Social discounting and intergenerational Pareto," *Econometrica*, 86, 1537–1567.

Fuster, L., A. Imrohoroglu, and S. Imrohoroglu (2007): "Elimination of social security in a dynastic framework," *The Review of Economic Studies*, 74, 113–145.

Galperti, S. and B. Strulovici (2017): "A theory of intergenerational altruism," *Econometrica*, 85, 1175–1218.

Gerlagh, R. and M. Liski (2017): "Consistent climate policies," *Journal of the European Economic Association*, 16, 1–44.

Gollier, C. and R. Zeckhauser (2005): "Aggregation of heterogeneous time preferences," *Journal of Political Economy*, 113, 878–896.

Gonzalez, F. M., I. Lazkano, and S. A. Smulders (2018): "Intergenerational altruism with future bias," *Journal of Economic Theory*, 178, 436–454.

Gul, F. (1991): "A theory of disappointment aversion," *Econometrica*, 59, 667–686.

Halevy, Y. (2008): "Strotz meets Allais: diminishing impatience and the certainty effect," *The American Economic Review*, 98, 1145–1162.

Hori, H. and S. Kanaya (1989): "Utility functionals with nonpaternalistic intergenerational altruism," *Journal of Economic Theory*, 49, 241–265.

Imrohoroglu, A. and K. Zhao (2018): "The Chinese saving rate: long-term care risks, family insurance, and demographics," *Journal of Monetary Economics*, 96, 33–52.

Iverson, T. and L. Karp (2020): "Carbon taxes and climate commitment with non-constant time preference," *The Review of Economic Studies*, 88, 764–799.

Jackson, M. O. and L. Yariv (2015): "Collective dynamic choice: the necessity of time inconsistency," *American Economic Journal: Microeconomics*, 7, 150–178.

Karp, L. (2016): "Discounting utility and the evaluation of climate policy," in *The economics of the global environment: catastrophic risks in theory and policy*, ed. by G. Chichilnisky and A. Rezai, Springer International Publishing, 423–446.

Koopmans, T. C. (1960): "Stationary ordinal utility and impatience," *Econometrica*, 28, 287–309.

KOPCZUK, W. AND J. P. LUPTON (2007): "To leave or not to leave: the distribution of bequest motives," *The Review of Economic Studies*, 74, 207–235.

KOTLIKOFF, L. AND L. SUMMERS (1981): "The role of intergenerational transfers in aggregate capital accumulation," *Journal of Political Economy*, 89, 706–32.

KREPS, D. M. AND E. L. PORTEUS (1978): "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica*, 46, 185–200.

LAIBSON, D. (1997): "Golden eggs and hyperbolic discounting," *Quarterly Journal of Economics*, 112, 443–77.

LAITNER, J. (1988): "Bequests, gifts, and social security," *The Review of Economic Studies*, 55, 275–299.

LE VAN, C. AND Y. VAILAKIS (2005): "Recursive utility and optimal growth with bounded or unbounded returns," *Journal of Economic Theory*, 123, 187–209.

LEININGER, W. (1986): "The existence of perfect equilibria in model of growth with altruism between generations," *Review of Economic Studies*, 53, 349–368.

LIZZERI, A. AND L. YARIV (2017): "Collective self-control," *American Economic Journal: Microeconomics*, 9, 213–44.

LOURY, G. C. (1981): "Intergenerational transfers and the distribution of earnings," *Econometrica*, 49, 843–867.

MAZZOCCO, M. (2007): "Household intertemporal behaviour: a collective characterization and a test of commitment," *The Review of Economic Studies*, 74, 857–895.

MERTENS, J.-F. AND T. PARTHASARATHY (1987): "Equilibria for discounted stochastic games," C.O.R.E. Discussion Paper 8750.

MILLNER, A. (2020): "Nondogmatic social discounting," *American Economic Review*, 110, 760–75.

MILLNER, A. AND G. HEAL (2018): "Time consistency and time invariance in collective intertemporal choice," *Journal of Economic Theory*, 176, 158 – 169.

MOOKHERJEE, D. AND S. NAPEL (2019): "Poverty dynamics and efficiency effects of conditional cash transfers," MS.

OBARA, I. AND J. PARK (2017): "Repeated games with general discounting," *Journal of Economic Theory*, 172, 348 – 375.

PAVONI, N. AND H. YAZICI (2016): "Intergenerational disagreement and optimal taxation of parental transfers," *The Review of Economic Studies*, 84, 1264–1305.

PEARCE, D. (2008): "Nonpaternalistic sympathy and the inefficiency of consistent intertemporal plans," in *Foundations in economic theory: a volume in honor of Hugo F. Sonnenschein*, ed. by M. O. Jackson and A. McLennan, Springer.

PELEG, B. AND M. E. YAARI (1973): "On the existence of a consistent course of action when tastes are changing," *Review of Economic Studies*, 40, 391–401.

PEROZEK, M. G. (1998): "A reexamination of the strategic bequest motive," *Journal of Political Economy*, 106, 423–445.

PHELAN, C. (2006): "Opportunity and social mobility," *The Review of Economic Studies*, 73, 487–504.

PHELAN, C. AND E. STACCHETTI (2001): "Sequantial equilibria in a Ramsey tax model," *Econometrica*, 69, 1491–1518.

PHELPS, E. AND R. POLLAK (1968): "On second best national savings and game equilibrium growth," *Review of Economic Studies*, 35, 195–199.

PIKETTY, T. AND G. ZUCMAN (2015): "Wealth and inheritance in the long run," in *Handbook of income distribution*, ed. by A. B. Atkinson and F. Bourguignon, Elsevier, vol. 2, 1303–1368.

POLLAK, R. A. (1968): "Consistent planning," *Review of Economic Studies*, 35, 201–208.

PRELEC, D. AND G. LOEWENSTEIN (1991): "Decision making over time and under Uuncertainty: a common approach," *Management Science*, 37, 770–786.

RAY, D. (1987): "Nonpaternalistic intergenerational altruism," *Journal of Economic Theory*, 41, 112 – 132.

RICHTER, M. (2021): "A welfare theorem for time-inconsistent discounters: patience convexity and modified Pareto efficiency," MS.

SAEZ-MARTI, M. AND J. WEIBULL (2005): "Discounting and altruism towards future decision-makers," *Journal of Economic Theory*, 122, 254–266.

SAITO, K. (2015): "Impure altruism and impure selfishness," *Journal of Economic Theory*, 158, 336–370.

SLEET, C. AND Ş. YELTEKIN (2016): "On the computation of value correspondences for dynamic games," *Operations Research*, 2, 174–186.

STROTZ, R. H. (1956): "Myopia and inconsistency in dynamic utility maximization," *Review of Economic Studies*, 23, 165–180.

YELTEKIN, S., Y. CAI, AND K. L. JUDD (2017): "Computing equilibria of dynamic games," *Operations Research*, 65, 337–356.

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