Antony Millner, Geoffrey Heal

Time consistency and time invariance in collective intertemporal choice

Article (Published version) (Refereed)

Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 176 (2018) 158–169

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Time consistency and time invariance in collective intertemporal choice

Antony Millner [a,*], Geoffrey Heal [b,c]

[a] *Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science, Houghton St, London, WC2A 2AE, UK*

[b] *Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027, USA*

[c] *NBER, USA*

Received 5 May 2017; final version received 19 February 2018; accepted 3 March 2018
Available online 6 March 2018

Abstract

Recent work on collective intertemporal choice suggests that non-dictatorial social preferences are generically time inconsistent. We argue that this claim conflates time consistency with two distinct properties of preferences: stationarity and time invariance. While time invariance and stationarity together imply time consistency, the converse does not hold. Although non-dictatorial social preferences cannot be stationary, they may be time consistent if time invariance is abandoned. If individuals are discounted utilitarians, revealed preference provides no guidance on whether social preferences should be time consistent or time invariant. Nevertheless, we argue that time invariant social preferences are often normatively and descriptively problematic.

JEL classification: D60; D71; D90

Keywords: Collective decisions; Intertemporal choice; Time consistency

1. Introduction

Many important decisions in economic life require groups of people with heterogeneous time preferences to implement a collective consumption plan. Examples abound: families must de-

* Corresponding author.
E-mail addresses: a.millner@lse.ac.uk (A. Millner), gmh1@columbia.edu (G. Heal).

https://doi.org/10.1016/j.jet.2018.03.002

cide on savings and intra-household resource allocation, partners in a firm must decide how to distribute profits between payouts to themselves and capital investments, communities with property rights over a natural resource must decide on an extraction plan, and resource rich countries must decide how to consume the proceeds from their sovereign wealth funds. In each of these examples an asset is held in common and is consumed dynamically over time, and the stake-holders in the decision very often have heterogeneous time preferences (Frederick et al., 2002). How should such decisions be evaluated and made, given people's different attitudes to time?

Recent theoretical and applied work (Jackson and Yariv, 2014, 2015; Adams et al., 2014) suggests that any attempt to make such collective intertemporal choices in a non-dictatorial fashion is doomed to confront a time inconsistency problem (see also earlier observations by Marglin, 1963; Feldstein, 1964; and Zuber, 2011). This paper argues that this finding is due to a conflation of time consistency with two *a priori* distinct properties of preferences: Stationarity and Time Invariance. Stationarity, introduced by Koopmans (1960), is an independence property of preferences, while time invariance requires preferences over future consumption streams to be the same in each evaluation period (this does not rule out preference reversals – quasi-hyperbolic time preferences, for example, are time invariant). Jackson and Yariv (2014, 2015), for example, show a necessary conflict between collective intertemporal choice and stationarity, but assume that social preferences must be time invariant. Yet while time invariance and stationarity together imply time consistency, the converse does not hold; time invariance is *not* required for time consistency. If the differences between these three properties of intertemporal preferences are made explicit it becomes clear that time consistent collective intertemporal choice *is* possible.

The differences between time consistent and time invariant social preferences are particularly subtle when individuals' preferences are discounted utilitarian, as is commonly assumed in the literature on the aggregation of time preferences. In this case, time consistency and time invariance are equivalent properties of individuals' preferences[1] (Halevy, 2015), but *not* of social preferences. We show that if individuals have heterogeneous discount factors, and social preferences are utilitarian and non-dictatorial, these preferences may be time consistent or time invariant, but not both. Since these properties are indistinguishable for discounted utilitarian individuals, but not for social preferences, the choice to impose one or the other property at the social level is ultimately a normative one – neither choice is more or less consistent with individuals' preferences. We argue however that time invariance is likely to be a problematic feature of social preferences – both normatively and descriptively – for within-group intertemporal choice. Time invariance may however be more descriptively plausible for between-group choice, as might occur in intergenerational decision-making. We relate these observations to the empirical literature that seeks to test the time consistency of household behaviour, suggesting that the definition of time consistency it employs is too restrictive.

2. Time consistency, time invariance, and stationarity

To make our argument we begin by providing definitions of three distinct properties of intertemporal preferences: time consistency, time invariance, and stationarity. Once we have defined these properties, it will immediately be clear that any two of them implies the third, as

[1] Discounted utilitarian preferences are also stationary.

observed in a simplified setting by Halevy (2015). We then show that the notion of time consistency used by e.g. Jackson and Yariv (2014, 2015); Adams et al. (2014) is in fact the conjunction of time invariance and stationarity. Time invariance and stationarity imply time consistency, but the converse does not hold.

We consider infinite consumption streams $\mathbf{x} = (x_\tau)_{\tau\in\mathbb{N}}$, where $\mathbb{N}$ is the set of non-negative integers, and x_τ takes values in some set A. The set of all such streams is denoted C^∞. Let h_τ be a history of consumption at evaluation period $\tau \geq 0$, and define H_τ to be the set of all possible histories at τ. Thus $h_\tau \in H_\tau := \{(x_0, \ldots, x_{\tau-1})\}$, and we define $H_0 = \emptyset$, where $\emptyset$ is the empty history. The set of all possible histories is denoted by $H = \cup_{\tau=0}^{\infty} H_\tau$, with generic element h. A *future* consumption stream $\mathbf{c}$ at a history h_τ is an infinite stream of *future* consumption values. Since the set of future consumption streams at any history h_τ is isomorphic to the set of initial consumption streams C^∞, we abuse notation slightly, and denote generic future consumption streams at any history h_τ as $\mathbf{c} = (c_0, c_1, \ldots) \in C^\infty$, where c_t denotes the consumption value realized t time steps in the future (i.e. at calendar time $\tau + t$).

Preferences are defined over future consumption streams $\mathbf{c}$ at each history $h \in H$. We assume a cardinal representation of preferences $V_h(\mathbf{c})$, and distinguish the following three properties of preferences:

Definition 1.

(a) *Time consistency*: For all $h \in H, a \in A, \mathbf{c}, \mathbf{c}' \in C^\infty$,

$$V_h((a, \mathbf{c})) \geq V_h((a, \mathbf{c}')) \iff V_{(h,a)}(\mathbf{c}) \geq V_{(h,a)}(\mathbf{c}') \tag{1}$$

(b) *Time invariance*: For all $h \in H, a \in A, \mathbf{c}, \mathbf{c}' \in C^\infty$,

$$V_h(\mathbf{c}) \geq V_h(\mathbf{c}') \iff V_{(h,a)}(\mathbf{c}) \geq V_{(h,a)}(\mathbf{c}') \tag{2}$$

(c) *Stationarity*: For all $h \in H, a \in A, \mathbf{c}, \mathbf{c}' \in C^\infty$,

$$V_h((a, \mathbf{c})) \geq V_h((a, \mathbf{c}')) \iff V_h(\mathbf{c}) \geq V_h(\mathbf{c}') \tag{3}$$

Time consistency is a consistency relationship between preferences at different histories. It rules out preference reversals, and implies that if an optimal plan is implemented today, it will remain optimal to follow it tomorrow. Time invariance also relates preferences at different histories, but is an altogether different property from time consistency. It requires that preferences be independent of translations of the time axis – shifting preferences forwards or backwards in time has no effect on rankings of future consumption streams. An immediate consequence of time invariance is that preferences are *history independent*.[2] Finally, stationarity is an independence property of preferences at each *fixed* history. This property is at the heart of Koopmans' (1960) axiomatization of discounted utilitarian time preferences. Koopmans was clearly aware that stationarity is a separate property from time invariance, and also has no necessary implications for the time consistency of preferences:

[2] Formally, preferences are history independent if for all times τ, all histories $h_\tau, h'_\tau \in H_\tau$, and all future consumption streams $\mathbf{c}, \mathbf{c}'$, $V_{h_\tau}(\mathbf{c}) \geq V_{h_\tau}(\mathbf{c}') \Rightarrow V_{h'_\tau}(\mathbf{c}) \geq V_{h'_\tau}(\mathbf{c}')$. All time invariant preferences are history independent, but not all history independent preferences are time invariant.

"[Stationarity] does *not* imply that, after one period has elapsed, the ordering then applicable to the 'then' future will be the same as that now applicable to the 'present' future. All postulates are concerned with only one ordering, namely that guiding decisions to be taken in the present." – Koopmans et al. (1964)

Halevy (2015) has observed in a simplified setting[3] that any two of the three properties in Definition 1 implies the third. Inspection of the definitions shows that this observation holds in general.

To make a connection between these definitions and the definition of time consistency commonly deployed in the literature (see e.g. Jackson and Yariv, 2014, 2015; Adams et al., 2014) in the simplest possible way, suppose that preferences are time separable, i.e.

$$V_h(\mathbf{c}) = \sum_{t=0}^{\infty} U_t^h(c_t) \tag{4}$$

where the sequence of utility functions $\{U_t^h(c_t)\}$ may be conditional on history h.[4] Under this restriction, the definition of time consistency employed by these authors reduces to the conjunction of stationarity and time invariance.[5]

The crux of the argument in Jackson and Yariv (2014, 2015) is a demonstration that non-dictatorial social preferences cannot be stationary if individuals have heterogeneous discount factors. The proof of this finding does not require much more than an application of the classic results of Koopmans (1960). Koopmans showed that stationarity requires the following constraints on the sequences of utility functions $\{U_t^h(c)\}$:

$$U_t^h(c) = U^h(c)(\beta_h)^t \tag{5}$$

[3] Halevy (2015) only considers preferences over timed consumption. Our definitions allow for preferences over full consumption streams.

[4] Separable preferences satisfy Koopmans' independence axiom, an additional assumption in Jackson and Yariv's (2015) definition of time consistency, which is in fact irrelevant to the question of time consistency in their case. Time consistency follows solely from the first of their stated conditions on preferences, i.e. the conjunction of stationarity and time invariance. Since the received wisdom in the literature, formalized by Jackson and Yariv (2015), is that non-dictatorial social preferences are *necessarily* time inconsistent, we need only refute this claim for a subset of preferences to make our point. Hence we lose nothing by focussing on a separable preferences, which are the focus of the empirical work in Jackson and Yariv (2014); Adams et al. (2014).

[5] Time invariance is implicit in the fact that the preferences Jackson and Yariv (2015) study do not carry an index that tells us the history of consumption, or the calendar time. Adams et al. (2014) make explicit reference to both stationarity and time invariance in their definition. Jackson and Yariv (2015) go some way towards acknowledging the role of time invariance in their definition of time consistency. They describe their 'stationarity' property informally as follows: 'Stationarity, in the sense that rankings of consumption streams do not depend on when they occur'. This description emphasizes time invariance, but is not an accurate description of their stated 'stationarity' property, which is *both* stationarity (an independence property of preferences at a fixed history) *and* time invariance (a relation between preferences at different histories). In footnote 18 of their paper they go on to say that: "There is a large literature that...views an agent as [time] consistent whenever plans of action are not overturned over time. Whenever consumption streams are evaluated in the same way in each period (so that agents do not have dated utility functions), the concepts are similar." What is not clear from this comment is how the main conclusion, i.e. the necessity of time inconsistency of collective intertemporal choice, relies critically on the time invariance property, and may be overturned if time invariance is abandoned. Moreover, even if individuals' preferences are not dated, as would occur if they are discounted utilitarians, identifying time consistency with the conjunction of stationarity and time invariance leads to constraints on social preferences that are very different to those that follow from the definition in (1).

for some history dependent utility function $U^h(c)$ and some discount factors $\beta_h \in [0, 1)$. Since only a single discount factor β_h is permitted at each history h in this representation, it is a simple matter to show that any non-dictatorial social preference cannot be stationary if individuals have preferences of the form (5), but their discount factors β_h are heterogeneous.[6] This follows since any such social preference must reflect the heterogeneity in individuals' preferences, which is not permitted by the stationary preferences (5).

To see why non-stationarity leads to a conflict with time consistency, note that we have already observed that:

(Time Consistency) AND (Time Invariance) $\Rightarrow$ Stationarity.

Taking the contrapositive of this implication,

NOT Stationary $\Rightarrow$ (NOT Time Invariant) OR (NOT Time Consistent).

Since non-dictatorial social preferences are not stationary, and the definition of time consistency in e.g. Jackson and Yariv (2014, 2015) assumes time invariance, a conflict with time consistency occurs. However, as is clear from the above implication, an alternative possibility exists. If we abandon time invariance of social preferences, the possibility of time consistency remains open.

3. Time consistent utilitarian social preferences

To demonstrate that non-dictatorial social preferences can indeed be time consistent, we follow Jackson and Yariv (2014, 2015) and assume that each individual i in a group G has discounted utilitarian time preferences:

$$V_h^i(\mathbf{c}) = \sum_{t=0}^{\infty} U^i(c_t)(\beta_i)^t \tag{6}$$

These preferences are stationary and time invariant, and hence time consistent. It follows that revealed preference cannot disentangle time consistency and time invariance for discounted utilitarian preferences.

While time consistency and time invariance are inseparable for discounted utilitarian agents, this is not the case for collective choices that are sensitive to the heterogeneity in agents' preferences, as we now demonstrate. We follow Jackson and Yariv (2014) and focus on utilitarian social preferences, and assume (without loss of generality, for our purposes) that consumption streams are common to all individuals. In addition, we assume that social preferences, like individuals' preferences, are time separable. Thus, we define social preferences at any history h_τ as:

$$\begin{aligned} W_{h_\tau}(\mathbf{c}) &= \sum_{i \in G} w^i(h_\tau) V_{h_\tau}^i(\mathbf{c}) \\ &= \sum_{i \in G} w^i(h_\tau) \sum_{t=0}^{\infty} U^i(c_t)(\beta_i)^t \end{aligned} \tag{7}$$

where $w^i(h_\tau)$ is the weight assigned to individual i when the history of consumption at calendar time τ is h_τ, and we assume that $w^i(h_\tau) \geq 0$, $\sum_{i \in G} w^i(h_\tau) = 1$ for all h_τ.

[6] See Proposition 1 in Jackson and Yariv (2015).

The following result gives conditions under which these preferences are non-dictatorial and time consistent/time invariant:

Proposition 1. *Assume that there exist indices i, j such that $w^i(h_\tau) > 0, w^j(h_\tau) > 0$ for all $h_\tau \in H$, and $\beta_i \neq \beta_j$. Then*

1. *Social preferences (7) are **time consistent** if and only if the welfare weights $w^i(h_\tau)$ are of the form:*

$$w^i(h_\tau) = \frac{w_0^i(\beta_i)^\tau}{\sum_{j\in G} w_0^j(\beta_j)^\tau}, \tag{8}$$

where $w_0^i \geq 0$, $\sum_{i\in G} w_0^i = 1$.

2. *Social preferences (7) are **time invariant** if and only if the welfare weights $w^i(h_\tau)$ are of the form:*

$$w^i(h_\tau) = y_i, \tag{9}$$

where $y_i \geq 0, \sum_{i\in G} y_i = 1$.

Proof. 1. The definition of time consistency in (1) requires

$$W_{h_\tau}((a,\mathbf{c})) \geq W_{h_\tau}((a,\mathbf{c}')) \iff W_{(h_\tau,a)}(\mathbf{c}) \geq W_{(h_\tau,a)}(\mathbf{c}') \tag{10}$$

for any $h_\tau, a, \mathbf{c}, \mathbf{c}'$. Substituting (7) into the left hand side of (10), it becomes

$$\sum_i w^i(h_\tau)U^i(a) + \sum_i \sum_{t=0}^{\infty} w^i(h_\tau)U^i(c_t)(\beta_i)^{t+1}$$

$$\geq \sum_i w^i(h_\tau)U^i(a) + \sum_i \sum_{t=0}^{\infty} w^i(h_\tau)U^i(c_t')(\beta_i)^{t+1}$$

$$\iff \sum_i \sum_{t=0}^{\infty} [w^i(h_\tau)\beta_i]U^i(c_t)(\beta_i)^t \geq \sum_i \sum_{t=0}^{\infty} [w^i(h_\tau)\beta_i]U^i(c_t')(\beta_i)^t$$

The right hand side of (10) is equivalent to

$$\sum_i \sum_{t=0}^{\infty} w^i((h_\tau,a))U^i(c_t)(\beta_i)^t \geq \sum_i \sum_{t=0}^{\infty} w^i((h_\tau,a))U^i(c_t')(\beta_i)^t$$

Comparing the expressions for the left and right hand sides of (10), we see that time consistency holds iff for all $h_\tau \in H$ and $a \in A$,

$$w^i((h_\tau,a)) \propto \beta_i w^i(h_\tau).$$

This can only occur if

$$w^i(h_\tau) = K w_0^i(\beta_i)^\tau$$

for some constants w_0^i, $K > 0$. Imposing the constraint $\sum_{i\in G} w^i(h_\tau) = 1$ yields the result.

2. Since time invariance requires $W_{h_\tau}(\mathbf{c})$ to be independent of τ, the result follows. □

Unlike the single agent case, time consistency and time invariance are not indistinguishable for utilitarian social preferences. The welfare weights (8) violate time invariance since with this choice preferences over future consumption streams vary with the evaluation date τ. Although individual agents have time invariant preferences, social preferences behave as if there is a fixed origin of the time axis at $\tau = 0$. Similarly, time invariant social preferences are not time consistent, as the welfare weights (9) are not of the form (8). In this case social preferences can exhibit reversals with the passage of calendar time, even though individuals' preferences cannot.

Observing that the conjunction of time consistency and time invariance implies stationarity, it is immediately clear from (5) that the only way it is possible to achieve both time consistency and time invariance in social preferences is if all individuals' utilities are discounted with the same discount factor β, and the welfare weights $w^i(\tau)$ are independent of τ:

$$\tilde{W}_{h_\tau}(\mathbf{c}) = \sum_{t=0}^{\infty}[\sum_i \tilde{w}^i U^i(c_t)]\beta^t. \tag{11}$$

This is the constraint on preferences used by Adams et al. (2014) in their empirical tests of the time consistency of observed choice behaviour. Proposition 1 shows that this is a substantially more stringent constraint than is required by time consistency. Time consistency does not rule out the possibility of social preferences reflecting the heterogeneity in individuals' discount factors, provided that the welfare weights attached to individuals' utilities evolve appropriately with calendar time.

We now make an observation that clarifies the possible interpretations of time consistent social preferences when agents have heterogeneous discount factors. The key point is that even if a discounted utilitarian agent's *preferences* over future consumption streams are known, there are still multiple ways of defining her *welfare* (see also Caplin and Leahy, 2004). To see this the following definitions will be useful:

Definition 2. Consider a discounted utilitarian agent with preferences (6). Let Q_τ be a measure of this agent's *welfare* at time τ. We say that the agent is:

(a) *Forward-looking* if $Q_\tau = \sum_{t=\tau}^{\infty} U(c_t)\beta^{t-\tau}$ for all τ
(b) Experiences *lifetime utility* if $Q_\tau = \left[\sum_{t=0}^{\tau-1} U(c_t)\beta^t\right] + \beta^\tau\left[\sum_{t=\tau}^{\infty} U(c_t)\beta^{t-\tau}\right]$ for all τ.

These two definitions of welfare are clearly equivalent from the perspective of the agent's preferences over future consumption streams. The next result shows that the fact that agents' welfare is underdetermined by their preferences has consequences for the interpretation of social preferences:

Lemma 1. *If utilitarian social preferences are time consistent and non-dictatorial, the following interpretations of (7) are equivalent:*

- *Individuals have forward-looking welfare measures, and social preferences have τ-dependent welfare weights (8).*
- *Individuals experience lifetime utility, and welfare weights are constant.*

Proof. Suppose that individuals experience lifetime utility, and that the welfare weights $w^i(h_\tau)$ are constant and equal to w_0^i. Then social preferences at any calendar time τ are given by:

$$W_\tau(\mathbf{c}) = \sum_i w_0^i \sum_{t=0}^{\tau-1} U(x_t)(\beta_i)^t + \sum_i w_0^i \sum_{t=0}^{\infty} U(c_t)(\beta_i)^{t+\tau} \tag{12}$$

where x_t are values of consumption in the past, and c_t are consumption values in the future. Since the x_t cannot be changed and have no influence on preferences over future consumption streams, these preferences are equivalent to

$$W_\tau(\vec{\mathbf{c}}) = \sum_i w_0^i (\beta_i)^\tau \sum_{t=0}^{\infty} U(c_t)(\beta_i)^t \tag{13}$$

which are in turn equivalent to time consistent social preferences with welfare weights (8). □

The ambiguity in the interpretation of social preferences is a consequence of the fact that we cannot tell whether individuals with discounted utilitarian preferences (6) have forward-looking welfare measures, or experience lifetime utility. Both interpretations give rise to the same individual rankings of future consumption streams, and are thus indistinguishable based on revealed preference. The lifetime utility interpretation of time consistent utilitarian preferences suggests that in any circumstance where cumulative outcomes over a period of time are seen as salient for the purposes of social comparisons, time consistency will be a compelling property, and time invariance may not.[7]

4. Time consistency or time invariance?

We have argued that there is no conflict between time consistency and non-dictatorial collective intertemporal choice, provided that time invariance is abandoned. Should we nevertheless insist that social preferences be time invariant, thus ruling out the possibility of time consistency? Given that revealed preference cannot tell us which property to adopt if individual preferences are discounted utilitarian, how should we make this choice? In order to demonstrate what is at stake when choosing to model social preferences as time invariant it is helpful to consider the following stylized scenario:

Ada and Bertha are identical twins; both are childless, single, and have identical wealth. On their 60th birthday they receive news that a distant cousin they've never met has died, and they will inherit his fortune. Their cousin died intestate, so there is no will to specify how the bequest should be divided between them. Ada is naturally impatient, while Bertha is patient. How should a utilitarian social planner allocate the bequest between them?

Suppose that the planner is time consistent. In this case first best intertemporal allocations can be decentralized by establishing property rights over the bequest, and then allowing each sister to consume her share as she pleases. Suppose that the planner's ethics dictate that an equal share of the bequest be given to each sister. The planner makes this allocation, and walks away happy in the knowledge that she need never revisit this decision.

[7] Similarly to Lemma 1, if social preferences are time invariant the following interpretations of (7) are equivalent:

- Individuals have forward-looking welfare measures, and social preferences have constant welfare weights (9).
- Individuals experience lifetime utility, and social preferences have τ-dependent welfare weights, $w^i(\tau) \propto y_i(\beta_i)^{-\tau}$.

We focus on the time consistent case, as the lifetime utility interpretation is more natural in this case.

Now consider a time invariant planner, and suppose that she also initially allocates equal shares to the sisters, before departing on other business, never intending to return. To her surprise however, the planner is asked to revisit the sisters 10 years later and assess their well-being. Ada has led a wild life in the intervening decade, consuming her inheritance much faster than Bertha. But the planner asks herself: what has changed since the sisters were 60? Although Ada has had a better life than Bertha, this is in the past and can't be changed. Since the circumstances are the same, except that the sisters happen to be 10 years older, the planner seizes some of Bertha's carefully saved cash, and reallocates it to Ada, so that their holdings of what remains of the bequest are again equal. Ada immediately spends her new wealth on fast cars and fancy footwear, while Bertha continues to put something away for a rainy day.

We believe that most people would intuitively feel that Ada should be made to take responsibility for her past excesses, and Bertha should not be penalized for her frugality in this scenario. Thus, many would believe that social preferences should be time consistent, not time invariant, in this case. A non-paternalistic reason for this could be that individuals experience 'memory utility' from past consumption (see e.g. Strotz, 1955; Caplin and Leahy, 2004), as suggested by the lifetime utility interpretation of time consistent social preferences in Lemma 1. Although individuals are not able to express this backward looking aspect of preferences directly through their choices, a social planner would be doing them a disservice if she neglected this for the purposes of social welfare computations. A purely normative reason for this could be that there is an evaluation date $\tau = 0$ that is ethically salient, even if individuals are in fact exclusively forward looking. In our scenario, the date of the initial bequest stands out as a natural 'origin' of the time axis. Many people would feel that cumulative utility from the date of the bequest onwards is ethically relevant to the task of apportioning the bequest.

Are there situations where time invariance might seem a more appropriate assumption than time consistency? We believe that modeling social preferences as time invariant might be more plausible if we are attempting to describe conflicts between successive groups of *different* people, as might occur in e.g. intergenerational decision-making. To illustrate, consider the following example:

A nation discovers an enormous oil deposit in its territorial waters. The deposit is large enough to generate rents for generations to come, which will be invested in public goods that benefit all citizens equally. Citizens are born in different time periods, and have finite lives.[8] Suppose that each citizen is either Avaricious (i.e. favours a high intergenerational discount rate), or Benevolent (i.e. favours a low intergenerational discount rate). In each period the currently living citizens must decide how much oil to extract. How is this likely to be done?

A 'pure' normative approach to this problem would require the construction of a social welfare function that aggregates the lifetime utilities of all individuals born in any time period, past, present, or future. To illustrate this procedure, define an indicator function:

$$\Delta_{i,s}(\tau) = \begin{cases} 1 & \text{if individual } i \text{ born at time } s \text{ is alive at time } \tau \\ 0 & \text{otherwise} \end{cases}$$

Using this definition, and letting $y^i(s) > 0$ be the welfare weight on individual i born in period s, such a normative welfare function might take the following form:

[8] Technically, the properties of stationarity and time invariance are only defined for infinite consumption streams, i.e. if agents are infinitely lived. Time consistency is however meaningful for finitely lived agents, and all three properties are meaningful for the social preferences (14) we define below. See Calvo and Obstfeld (1988) for a model in which agents have finite, but uncertain, lifetimes, which *de facto* induces preferences over infinite consumption streams.

$$W_{h_\tau}(\mathbf{c}) = \sum_{s=0}^{\tau-1}\sum_{i\in G} y^i(s)\left(\sum_{t=0}^{\infty}\tilde{U}^i(c_t)(\delta_i)^{t+\tau-s}\Delta_{i,s}(\tau+t)\right) + \sum_{s=\tau}^{\infty}\sum_{i\in G} y^i(s)\left(\sum_{t=0}^{\infty}\tilde{U}^i(c_{s-\tau+t})(\delta_i)^t\Delta_{i,s}(s+t)\right) \tag{14}$$

The first term in this expression captures the remaining lifetime utilities of individuals born at times $s < \tau$ who are still alive in the current evaluation period τ. The second term captures the lifetime utilities of those who are born at times $s \geq \tau$. Agents' views on the appropriate degree of social altruism towards future generations are irrelevant for the purposes of constructing the social welfare function (14), and may be discarded.[9] The discount factors δ_i that appear in this expression are agents' *private* discount factors on their own utilities $\tilde{U}^i(c)$, and have no necessary relationship to the *social* welfare weights $y^i(s)$, which are purely normative quantities. Social preferences constructed in this manner will be time consistent, as follows from an adaptation of Proposition 1 and Lemma 1 to this overlapping lifetimes setting. While this procedure yields a normative social preference relation that does not conflict with time consistency, from a descriptive perspective it makes strong demands on current living agents. In particular, the current generation is required to discard their own views on how much importance to give to future generations when deciding on their extraction policy.

What seems more likely to occur from a descriptive perspective is that the agents currently alive will decide how much oil to extract through a collective decision-making process that trades off the different degrees of altruism expressed by *current* individuals. If this process is efficient the preferences of generation τ may be represented by a welfare function of the form:

$$W_{h_\tau}(\mathbf{c}) = \sum_{t=0}^{\infty}\sum_{i\in G} w^i U^i(c_t)\beta_i^t, \tag{15}$$

where w^i is the weight given to agent i's opinion. The interpretation of the inputs to this welfare function is entirely different to that in (14). Here the utility functions $U^i(c)$ reflect i's opinion about *social* utility, rather than her private utility, and the discount factors β_i reflect her opinion on the appropriate degree of intertemporal *social* impatience. Since the decision-making process that determines current extraction will unfold in each generation, time invariance seems a reasonable property to impose if we are attempting to model the conflict between the preferences of successive groups of decision-makers.[10]

The crucial difference between the 'twins' and 'oil' scenarios we have sketched is that in the case of the twins we were concerned with intertemporal distribution *within* a fixed group of people over time, whereas in the oil case we were concerned with intertemporal distribution *between* different groups of people. In the former case all agents who are affected by a decision are present at all points of time, whereas in the latter only a subset of agents affected by a decision are present at each point in time. Time consistency seems a much more plausible descriptive property of social preferences than time invariance in the former case, since the ahistorical nature of time invariant preferences would be seen as deeply unfair by current agents. In the case of between-group choice however, although time invariance may do violence to the preferences of

[9] Note however that this is a paternalistic move if individuals experience anticipatory utility from the wellbeing they expect their descendants to achieve.

[10] Note that the composition of successive groups must be stable if the preferences (15) are to be time invariant.

previous or successive groups, these individuals have no voice in *current* decision processes. As a descriptive matter, it seems more likely that the interests of past and future groups are only relevant insofar as *current* decision-makers account for them. To our minds this makes time invariance a more plausible descriptive modeling assumption in this case. Of course, this argument makes no claims about the normative status of time invariance as a property of social welfare criteria.

5. Implications for the empirical literature

While we find time invariance to be a not implausible descriptive property in between-group choice, the empirical literature focusses exclusively on within-group choice. Our hypothesis that time consistent, but non-invariant, social preferences are likely to prevail in this case receives support from the innovative study of household consumption behaviour in Adams et al. (2014). Although they strongly reject the hypothesis that household behaviour can be described by time invariant and time consistent preferences of the form (11), they find considerable support for the hypothesis that household consumption behaviour can be rationalized by time consistent social preferences with time varying welfare weights (8). Adams et al. (2014) refer to this as the 'full efficiency' model, and claim that this identifies individual discount rate heterogeneity as a source of household time inconsistency. Yet the appropriate interpretation of these findings is that collective choices *can* often be rationalized by a time consistent utilitarian model, provided that household preferences account for the cumulative lifetime utility experienced by each household member.

In related work, Jackson and Yariv (2014) study a laboratory simulation of collective intertemporal choice in which subjects are asked to choose between collective consumption allocations on behalf of a group of individuals with heterogeneous, exogenously imposed, discount factors. They elicit preferences over pairs of collective consumption streams such as

$$\mathbf{C} = (105, 0, 0) \qquad \mathbf{C}' = (0, 160, 0)$$
$$\mathbf{C}'' = (0, 105, 0) \qquad \mathbf{C}''' = (0, 0, 160).$$

Stationarity requires the planner's preferences to obey $\mathbf{C} \succeq \mathbf{C}' \iff \mathbf{C}'' \succeq \mathbf{C}'''$. Jackson and Yariv (2014) observe almost unanimous violation of this property, leading them to conclude that their subjects' collective preferences are time inconsistent. Yet these choice tasks provide little information about the time consistency of planners' preferences. As (1) makes clear, a test of time consistency requires the construction of dynamic choice tasks, in which subject planners choose between a pair of collective consumption streams, the group is given the first period outcome of the planner's preferred stream, and the planner is then asked to choose between the continuations of the initial consumption streams.[11] No such dynamic choice tasks were considered in the Jackson and Yariv (2014) study – subjects always made once-off choices between *fixed* consumption streams. All we can say for sure is that the study provides strong evidence against stationarity, as should be expected when planners attempt to account for the heterogeneous preferences of group members.

[11] See Halevy (2015) for an experimental design that allows time consistency, time invariance, and stationarity to be disentangled.

Acknowledgments

Millner acknowledges support from the Centre for Climate Change Economics and Policy, funded by the UK Economic and Social Research Council, grant number ES/K006576/1. We are grateful to Geir Asheim, Marc Fleurbaey, Itzhak Gilboa, and two anonymous referees for very helpful comments.

References

Adams, A., Cherchye, L., Rock, B.D., Verriest, E., 2014. Consume now or later? Time inconsistency, collective choice, and revealed preference. Am. Econ. Rev. 104 (12), 4147–4183.

Calvo, G.A., Obstfeld, M., 1988. Optimal time-consistent fiscal policy with finite lifetimes. Econometrica 56 (2), 411–432.

Caplin, A., Leahy, J., 2004. The social discount rate. J. Polit. Econ. 112 (6), 1257–1268.

Feldstein, M.S., 1964. The social time preference discount rate in cost benefit analysis. Econ. J. 74 (294), 360.

Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time discounting and time preference: a critical review. J. Econ. Lit. 40 (2), 351–401.

Halevy, Y., 2015. Time consistency: stationarity and time invariance. Econometrica 83 (1), 335–352.

Jackson, M.O., Yariv, L., 2014. Present bias and collective dynamic choice in the lab. Am. Econ. Rev. 104 (12), 4184–4204.

Jackson, M.O., Yariv, L., 2015. Collective dynamic choice: the necessity of time inconsistency. Am. Econ. J. Microecon. 7 (4), 150–178.

Koopmans, T.C., 1960. Stationary ordinal utility and impatience. Econometrica 28 (2), 287–309.

Koopmans, T.C., Diamond, P.A., Williamson, R.E., 1964. Stationary utility and time perspective. Econometrica 32 (1/2), 82–100.

Marglin, S.A., 1963. The social rate of discount and the optimal rate of investment. Q. J. Econ. 77 (1), 95–111.

Strotz, R.H., 1955. Myopia and inconsistency in dynamic utility maximization. Rev. Econ. Stud. 23 (3), 165.

Zuber, S., 2011. The aggregation of preferences: can we ignore the past? Theory Decis. 70 (3), 367–384.