Modelling change in individual characteristics: an axiomatic framework

Franz Dietrich[1]

CNRS, France & University of East Anglia, U.K.

Abstract

Economic models describe individuals by underlying characteristics, such as the degree to which they like music, have sympathy, want success, need recognition, etc. In reality, such characteristics change through experiences: taste for Mozart changes through attending concerts, sympathy through meeting people, etc. Models typically ignore change, partly because it is unclear how to incorporate it. I develop a general axiomatic framework for defining, analysing and comparing rival models of change. Seemingly basic postulates on modelling change have strong implications, like irrelevance of the order in which someone has his experiences and 'linearity' of change. This is a step towards placing the modelling of change on axiomatic grounds and enabling non-arbitrary incorporation of change into economic models.

JEL Classification Codes: C0, C7, D0, D8.

Keywords: change, taste, sympathy, altruism, rational choice, axiomatic foundations, dynamic (in)consistency, externalities

1 Introduction

In much of economic modelling practice, nothing about an individual (except perhaps his information state) is taken to change over time. For instance, an individual engaged in a dynamic decision problem or game with stages $t = 1, 2, ..., T$ (T finite or infinite) is often assumed to maximise (the expectation of) a discounted sum $\sum_{t=1}^{T} \delta^t u(a_t)$, in which a_t is the period-t outcome (e.g. his period-t consumption bundle, or in the case of a game the period-t action profile), δ is

---

a discount factor, and $u(.)$ is the individual's intra-period utility function which, importantly, does not change (exogenously) with time t or (endogenously) with the outcomes of past periods. Such a preference specification precludes that the individual's period-t ability to enjoy the period's outcome depends on time or past outcomes. Gary Becker (1996) and many others stress the unrealistic nature of such an assumption: in real life, the pleasure derived from listening to classical music, consuming drugs, meeting friends, and so on, depends on time (kids differ from adults) and on the past consumption pattern (enjoyment of Bordeaux wine has to be learnt). Sen (1977, 1979, 1985), Rabin (1998) and many others stress the possibility to have, develop or lose other-regarding feelings that reflect sympathy, hate, reciprocity, identification or other attitudes: in a real-life repeated interaction within a couple, Ann may have changing feelings for Peter (depending on age and past events), where the state of these feelings in a period t determines how much pleasure Ann then receives from an outcome a_t that benefits Peter. All these scenarios are excluded by defining period-t utility invariably as $u(a_t)$. The (equilibrium) behaviour one derives in a decision problem or game would be more realistic if change in agent characteristics were successfully incorporated into agent preference.

I develop a unified axiomatic framework in which to define, analyse and compare rival models of change in a characteristic, such as taste for wine, identification with one's partner, risk aversion, alcohol addiction, need for recognition, impatience as identified with a discounting rate, personal or social capital as in Becker's consumer theory, and so on. More precisely, I model the individual as being at any moment in some *state s*, a real number that measures the characteristic of interest. The state changes under what I call *experiences*, where this term is used in its broadest and most flexible sense, covering both internal events (e.g. the effect of a drug, coming into puberty, getting Alzheimer) and external events (e.g. the smile of a child, an earthquake), and covering events either under the agent's own control (e.g. his moves in decision problem or game) or not under his control (e.g. moves of others or nature). The framework can be applied in several ways, for instance to study 'non radical' preference change which preserves stable fundamental preferences over fully specified worlds or world histories,[2] or to study fundamental preference change (dynamic inconsistency) by letting these preferences depend on a changing characteristic (e.g. an addiction level), or to study change in preference-unrelated characteristics such as beliefs. The present axiomatic approach responds to the discrepancy between recognized importance of change and lack of theoretical understanding of it.

The axioms on change introduced below have some flavour of 'rationality'

---

[2] For instance, in the above example's intertemporal utility $\sum_{t=1}^{T} \delta^t u(a_t)$ one could replace each intratemporal utility term $u(a_t)$ by a term $u(a_t, s)$ that depends on the agent's current state s (e.g. his taste for wine or feelings for someone), which can in turn be specified as a function of time t and/or of past outcomes $a_0, ..., a_{t-1}$.

postulates, though it is clear that principles of rationality alone cannot tell us how exactly someone's feelings or taste for wine will (should?) change in the face of certain experiences. I introduce some postulates, whose combination turns out to severely constrain and discipline change, forcing it to take a simple and convenient form potentially suitable for modelling practice. Specifically, the order in which experiences are made is irrelevant to overall change, and (adding further conditions) change is 'linear'. As I shall emphasise at different points, these findings can be interpreted either as providing welcome axiomatic support for modelling change in a simple way, or, by contraposition, as informing us that any change pattern without these simple features (of order-insensitivity or linearity) can only be modelled by violating basic axioms on change.

Mathematically, the key insight is that experiences can be viewed as *operators* (operating on individual constitutions) that can be *composed* (representing repeated experiences) and *ordered* (in terms of strength of experience). This allows me to apply basic theorems of ordered group theory and topological algebra by Hölder (1901), Huntington (1902), Arzél (1948), Tamari (1949), Alimov (1950) and Nakada (1951).

While the axiomatic approach to individual change is new, this paper is related, at least in its motivation, to a growing and diverse literature on endogeneity, i.e. on the dependence of human tastes and other characteristics on the environment, institutions, characteristics of others, and so on. This literature has added significantly to our understanding and offers concrete models incorporating endogeneity. See for instance Polak (1976), Bowles (1998) and Rabin (1998); on endogenous other-regarding feelings (such as sympathy, spitefulness, reciprocal feelings), see Rabin (1993), Fehr and Gächter (1998), Bolton and Ockenfels (2000), Sethi and Somanathan (2001) and (2003), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006), Dietrich (2008) and Maccheroni *et al.* (forthcoming); on endogenously changing fundamental preferences (i.e. dynamic inconsistency), see Strotz (1955-56), Hammond (1976), and O'Donoghue and Rabin (1999). Two notable approaches to modelling preference change are preference evolution models (e.g., Dekel et al. 2007) and the theory of adaptive utility (initiated by Cyert and DeGroot 1975). Each of these models does of course in its own particular way address change of the endogenous characteristic in question. But these models are not (and were not intended as) full-fledged, general 'change models', for a variety of reasons.[3] This paper aims for a full-fledged change model. The paper can also be viewed as a response to the non-unified character of our current theory of endogeneity, which indeed appears more as a disjunction of several special theories, each one designed for a particular human characteristic, environment or experimental setup.

---

[3] Notably, they usually do not give a proper formal account of the 'experiences' causing changes and of the 'revision rule'. So, they are lagging behind when compared to, say, models of belief revision or learning.

3

2 Change models

I now define a general notion of 'change model'. The notion is flexible: it abstracts from the particular characteristic of interest, so as to be applicable – at least in principle – to (change in) various kinds of characteristics, such as taste for goods, altruism, risk aversion, drug addiction, need for recognition, impatience (discounting rate), or personal or social capital. Formally, a *state space* is an arbitrary set $\mathbf{S}$, whose members represent the agent's possible states w.r.t. the relevant characteristic. Typically, a state is a number (e.g. a sympathy level or a parameter measuring taste for wine and entering the agent's Cobb-Douglas utility function over consumption bundles); or, it is a vector of numbers (e.g. the entire parameter vector of the individual's Cobb-Douglas utility function). If the state space is a subset of some Euclidean space $\mathbb{R}^n$, where it contains more than one state and is compact and convex, then it is called *Euclidean*, or more precisely *n-dimensional*. In particular, a *one-dimensional* state space takes the compact and convex form $\mathbf{S} = [a, b] \subseteq \mathbb{R}$ where $a < b$.

How does the state change? Let us first take a simple approach. By a *simple change model* (for a state space $\mathbf{S}$) I mean a triple $(\mathbf{S}, \mathbf{E}, (.|.))$ consisting of:

- the state space $\mathbf{S}$,
- an arbitrary set $\mathbf{E}$ (of possible *experiences*),
- a function $(.|.) : \mathbf{S} \times \mathbf{E} \rightarrow \mathbf{S}$ (the *revision rule*, which maps each pair of an initial state s and an experience e to the agent's new state $s|e$).

The set $\mathbf{E}$ contains all experiences the individual may have. As explained earlier, the term 'experience' is taken in its broadest sense as being *any* relevant influence on the individual. The mathematical structure of experiences is entirely general: experiences could be numbers, vectors, functions, elements of a metric space, or whatever the modeller wishes.

Simple change models give us a first, base-line approach to representing change. As an important example, Bayesian belief revision is a simple change model: states are probability functions, experiences are observed events, and revision is Bayesian updating. But we mainly focus on preference- rather than belief-related characteristics; there, simple change models are less natural.

But often this approach is 'too simple' to work. When attempting to define the revision rule, one may encounter a severe problem: the new state to which a given experience e leads is *underdetermined* by the old state s. So, s fails to encode enough information to determine the new state, i.e. to adequately define $s|e$. As an example, suppose states are altruism levels. The agent currently is in a state of high altruism s, but after a 'negative life experience' e acquires a state of low altruism s_l; so, $s|e = s_l$. Now suppose alternatively that, before having this experience, the agent has a 'positive life experience' e' which confirms the agent in his high altruism, thus leaving his state unchanged; so, $s|e' = s$. Though leaving the state unaffected, the interim experience has entrenched altruism more

4

deeply in the agent's psychology (a fact that the state fails to capture). As a result, the following 'negative life experience' e has less of an effect on the agent's state s, changing it merely to a state of medium altruism s_m; so, $s|e = s_m$. This is an immediate contradiction, since we cannot define $s|e$ both as s_l and as s_m. The lesson is that it may be unclear how to define revision since the effect of an experience on the state may depend on psychological information not contained in that state, such as information about 'state entrenchment'.

In response, I introduce a richer description of the agent than his state, to be called his 'constitution'. It can encode additional information (for instance about 'state entrenchment') which goes beyond the characteristic of interest, e.g. beyond the parameters of the agent's utility function over consumption bundles. Although it is the state, not the entire constitution, which ultimately matters, the constitution will play a key role: revision can now be defined on the level of constitutions, and change in state can be explained by change in underlying constitution. Formally, a *change model* (for a state space $\mathbf{S}$) is a tuple $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ consisting of:

- the state space $\mathbf{S}$,
- an arbitrary set $\mathbf{E}$ (of possible *experiences*),
- an arbitrary set $\mathbf{C}$ (of possible *constitutions*),
- a surjective function $(\bar{\cdot}) : \mathbf{C} \to \mathbf{S}$ (the *state projection*, which maps each constitution c to the agent's state $\bar{c}$ in that constitution),
- a function $(.|.) : \mathbf{C} \times \mathbf{E} \to \mathbf{C}$ (the *revision rule*, which maps each pair of an initial constitution c and an experience e to the agent's new constitution $c|e$),

such that two technical conventions to be defined shortly (namely (1) and (2)) are respected. The state projection extracts the ultimately relevant information from the agent's constitution. (The surjectivity assumption is there to ensure that $\mathbf{S}$ contains no impossible states.) The revision rule $(.|.)$ operates on the level of constitutions, not states. As the agent's constitution changes from c to $c|e$, his *state* changes from $\bar{c}$ to $\overline{c|e}$. A constitution c must indirectly encode not just the current state (namely, $\bar{c}$) but also the way the state reacts to future experiences. Does this imply that constitutions must be defined as highly complex objects, as complex as a genetic code, and too complex for practical purposes? This question will be of central interest to us (and I hope to bring some positive news). Mathematically, constitutions can be arbitrary objects (e.g. vectors), just as experiences.

This notion of a change model strictly generalizes that of a simple change model. Indeed, a simple change model $(\mathbf{S}, \mathbf{E}, (.|.))$ can be viewed as a special change model in which constitutions are identified with states, i.e. in which $\mathbf{C} := \mathbf{S}$ and $\bar{c} := c$ for all $c \in \mathbf{C}$.[4]

---

[4]In the so-defined change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$, the components '$\mathbf{C}$' and '$(\bar{\cdot})$' are of course redundant (so that the triple $(\mathbf{S}, \mathbf{E}, (.|.))$ suffices to describe the model). To be entirely precise,

The need to go beyond simple change models and work with general ones is not just an artifact of a mis-specified notion of 'state', as one might at first suspect. One might suspect that as soon as states are specified more richly by including all necessary information, one *can* define revision satisfactorily at the level of states without invoking a separate notion of 'constitution'. This impression is misguided. Far from allowing us to stick to a simple change model, this strategy of 'enriching states' leads directly to a general change model (albeit under different terminology). The reason is, firstly, that such 'informationally enriched states' are simply 'constitutions' in our initial sense, so that $\mathbf{S}$ collapses into $\mathbf{C}$. Secondly, since the 'enriched state' contains economically irrelevant information, one needs a function which extracts the relevant information from it; but the extracted information is simply the 'state' in our initial, non-enriched sense, and the function is the state projection $(\bar{\cdot})$. We are thus led back to a change model in our general, not our simple sense.

Change models obviously involve additional theoretical constructs, namely experiences and constitutions, which – like the traditional constructs of subjective probabilities and utilities – are hard or impossible to observe directly. What matters for testability and falsifiability purposes is that these constructs have observable implications, which they do in may applications.

Notation and definitions. I often drop brackets when it is clear how to place them; e.g. $c|e|e'$ stands for $(c|e)|e'$, and $c|e_1 \cdots |e_n$ stands for $(\cdots (c|e_1) \cdots)|e_n$ (interpreted as c if $n = 0$). I call the change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ a *submodel* of another one $(\hat{\mathbf{S}}, \hat{\mathbf{E}}, \hat{\mathbf{C}}, (\hat{\cdot}), (.|.))$ (and the latter a *supermodel* or *extension* of the former) if $\mathbf{S} \subseteq \hat{\mathbf{S}}$, $\mathbf{E} \subseteq \hat{\mathbf{E}}$, $\mathbf{C} \subseteq \hat{\mathbf{C}}$, $(\bar{\cdot}) = (\hat{\cdot})|_{\mathbf{C}}$ and $(.|.) = (\hat{.|.})\big|_{\mathbf{C} \times \mathbf{E}}$.

The two conventions. Call experiences e, e' *essentially identical* if they have the same effect on each constitution, i.e. if $c|e = c|e'$ for all constitutions c. By convention, the model describes experiences only as far as relevant:[5]

$$\text{no distinct experiences } e, e' \in \mathbf{E} \text{ are essentially identical.} \tag{1}$$

So, if losing a friend on Monday and doing so on Tuesday affect the individual in the same way, the two will be modelled as the same experience $e \in \mathbf{E}$ of 'losing a friend'. Hence, each $e \in \mathbf{E}$ in a sense represents an experience *type*, which makes it meaningful to experience e repeatedly.

What matters about a constitution c are the present and future states. Accordingly, I call constitutions c, c' *essentially identical* if $\overline{c|e_1 \cdots |e_n} = \overline{c'|e_1 \cdots |e_n}$ for all experience sequences $(e_1, ..., e_n)$ of any (possibly zero) length $n \geq 0$ (i.e. if $\bar{c} = \bar{c'}$ *and* $\overline{c|e} = \overline{c'|e}$ for all experiences e *and* $\overline{c|e_1|e_2} = \overline{c'|e_1|e_2}$ for all experiences e_1, e_2, etc.). By convention, the model describes constitutions only as far as

---

we have defined a proper change model only if the conventions (1) and (2) are respected.

[5]So, experiences may be formally identified with constitution transformations $\mathbf{C} \to \mathbf{C}$.

relevant:

$$\text{no distinct constitutions } c, c' \in \mathbf{C} \text{ are essentially identical.} \tag{2}$$

The conventions (1) and (2) impose no loss of generality.[6]

3   Examples and applications

I now give two formal examples of change models, followed by two concrete applications. (Later in Section 7 I give two further applications which, unlike the present ones, are partly geared to dynamic inconsistency.)

Example 1: the linear model. This is a particularly important change model, later shown to have several salient properties. For a given Euclidean state space $\mathbf{S} \subseteq \mathbb{R}^n$ (where $n \geq 1$), the *linear* change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ is defined by:

- the set of experiences $\mathbf{E} = \mathbf{S} \times (0, \infty)$. An experience $(s, x) \in \mathbf{E}$ is written s_x, with x interpreted as strength of experience and s as the state to which the individual is attracted. (If states are levels of sympathy for kids, the experience s_x of meeting a friendly kid might have high s.)
- the set of constitutions $\mathbf{C} = \mathbf{S} \times [0, \infty)$. A constitution $(s, x) \in \mathbf{C}$ is again written s_x, with s interpreted as the current state and x as strength of constitution, i.e. immunity to experience.
- the state projection given by $\overline{s_x} = s$ for all $s_x \in \mathbf{C}$. (So the 's' in a constitution s_x stands indeed for the current state.)
- the revision rule given by $s_x | \tilde{s}_{\tilde{x}} = \left(\frac{x}{x+\tilde{x}} s + \frac{\tilde{x}}{x+\tilde{x}} \tilde{s} \right)_{x+\tilde{x}}$ for all $s_x \in \mathbf{C}$ and $\tilde{s}_{\tilde{x}} \in \mathbf{E}$.

So, having an experience $\tilde{s}_{\tilde{x}}$ in a constitution s_x leads the state to change from s to a weighted average of s and $\tilde{s}$, with weights being determined by the strengths of the experience and the old constitution. The new constitution has strength $x + \tilde{x}$. So, the stronger the old constitution and the experience, the stronger the new constitution, and hence, the smaller the effect of future experience (which seems plausible in that future experience must then compete with a higher stock of past influences). Repeatedly applying the linear revision rule, the effect of an entire sequence of experiences $s^1_{x^1}, ..., s^t_{x^t}$ ($t \geq 0$) on a constitution s_x is given by

$$s_x | s^1_{x^1} \cdots | s^t_{x^t} = \left(\frac{xs + x^1 s^1 + ... + x^t s^t}{x + x^1 + ... + x^t} \right)_{x+x^1+...+x^t}. \tag{3}$$

Example 2: a non-parametric model. As a more abstract example, consider a change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ defined by:

- a one-dimensional state space $\mathbf{S} = [a, b] \subseteq \mathbb{R}$.

---

[6]If the conventions are initially violated, simply identify any essentially identical constitutions and identify any experiences $e, e' \in \mathbf{E}$ such that $c|e$ and $c|e'$ are essentially identical for all $c \in \mathbf{C}$.

- the set of constitutions $\mathbf{C}$ consisting of all finite measures on (the Borel-measurable sets of) $\mathbf{S}$ with a non-negative density function $f : \mathbf{S} \to [0, \infty)$. Here, $f(s)$ represents how much the individual currently 'tends' to state s.
- the set of experiences $\mathbf{E}$ consisting of all finite measures on $\mathbf{S}$ with a positive density function $f : \mathbf{S} \to (0, \infty)$. So, experiences are again measures, this time capturing tendencies in (the effect of) experience.
- the state projection given by $\bar{c} = Median(c)$ for all $c \in \mathbf{C}$.[7] The agent's state $\bar{c}$ represents a 'summary' or 'compromise' of all tendencies in his current constitution c.[8]
- the revision rule given by $c|e = c + e$ for all $c \in \mathbf{C}$ and $e \in \mathbf{E}$.

So, the agent's post-experience constitution is the sum of his old constitution c and the experience e; his state changes from $Median(c)$ to $Median(c + e)$.

I now sketch two applications, namely a decision-theoretic one followed by a game-theoretic one.

Application 1: the state as the taste of a Becker-type consumer. Consider an agent in an intertemporal consumption problem with T periods and K goods ($K, T \in \{1, 2, ...\}$). A (consumption) bundle is a vector $b = (b^1, ..., b^K) \in [0, \infty)^K$, with b^k denoting quantity of good k. A *(consumption) path* is a tuple $(b_t)_{t=1,...,T}$ of bundles, with b_t denoting the bundle consumed in period t. Following Becker (1996), taste for many goods (e.g. wine or classical music) depends on past consumption. Accordingly, let the individual's state s be a measure of his taste for goods. If in a period t the agent consumes bundle b with taste (state) s, he receives utility $u(b; s)$ in that period. The analytic form of $u(b; s)$ might belong to one of the classic parametric families (Cobb-Douglas, CES, ...), with s being one of the parameters or a vector of some or all of the parameters.[9]

Becker's important insight is that the past consumption pattern $b_1, ..., b_{t-1}$ affects period-t taste s; let us write $s = s(b_1, ..., b_{t-1})$ to capture the dependence. But Becker's theory gives no clear answers to our question of *how* taste changes, i.e. how we should specify $s(b_1, ..., b_{t-1})$ as a function of $b_1, ..., b_{t-1}$. This question matters notably for the intertemporal consumer problem of maximising intertemporal utility

$$U((b_t)_{t=1,...,T}) := \sum_{t=1}^{T} \delta^t u(b_t; s(b_1, ..., b_{t-1})) \tag{4}$$

---

[7]By definition, the median $m = Median(c)$ has the property that $c(\{s \in \mathbf{S} : s \leq m\}) = c(\{s \in \mathbf{S} : s \geq m\})$; if more than one $m \in \mathbf{S}$ has this property, $Median(c)$ is by convention the middle of the interval of all these m's (also other conventions would work).

[8]The median is a compromise in that it minimises the average distance to states (relative to the measure).

[9]Using the CES utility function, $u(b; s) = ((s_1 b^1)^\rho + ... + (s_K b^K)^\rho)^{1/\rho}$ for parameters $s_1, ..., s_K, \rho > 0$. The state s could be s_1 (taste for good 1), with the other parameters being fixed, i.e. unchanged. Or, s could be the vector of all parameters, a multi-dimensional state.

over consumption paths $(b_1, ..., b_T) \in [0, \infty)^{K \times T}$ subject to a budget constraint (with $\delta > 0$ denoting a fixed discounting factor).

If we for example adopt the linear change model, taste changes as follows. Constitutions and experiences are represented by strength-indexed states. Initially, the agent is in some constitution $c \equiv s_x$, i.e. his taste is s, entrenched to degree x. The experience of consuming a bundle b is identified with a strength-indexed state: $b \equiv s_{x^b}^b$; i.e., taste is attracted to s^b with strength x^b. Applying the linear revision rule, the new constitution is $\overline{c|b} = \frac{xs + x^b s^b}{x + x^b}$. More generally, applying (3), consumption of $b_1, ..., b_{t-1}$ leads to the new period-t taste given by:

$$s(b_1, ..., b_{t-1}) = \overline{c|b_1 \cdots |b_{t-1}} = \frac{xs + x^{b_1} s^{b_1} + ... + x^{b_{t-1}} s^{b_{t-1}}}{x + x^{b_1} + ... + x^{b_{t-1}}}.$$

If, by contrast, we follow Example 2's non-parametric change model (assuming that the state is one-dimensional), then taste changes rather differently, leading to different optimal consumption paths. The agent's initial constitution is represented by a measure c over $\mathbf{S}$, whose median defines the initial taste. The experience of consuming a bundle b is also identified with a measure $b \equiv \mu^b$. A consumption path $b_1, ..., b_{t-1}$ leads to the new period-t constitution $c|b_1 \cdots |b_{t-1} = c + \mu^{b_1} + ... + \mu^{b_{t-1}}$, whose median defines the new taste $s(b_1, ..., b_{t-1})$.

Application 2: state as the level of sympathy for another player. One often observes cooperative behaviour in repeated interactions with a prisoners'-dilemma-type monetary payoff structure. Arguably, such phenomena are often best explained not by postulating irrationality of entirely self-interested agents, and also not by postulating stable levels of sympathy (other-regardingness), but rather by allowing endogenous change in sympathy levels in response to the 'treatment' the player receives by others. Indeed, the level of sympathy for other people plausibly changes with their behaviour. We therefore have to model change in sympathy. As an illustration, consider a dynamic game with two players 1 and 2, perfect information, and stages $t = 0, 1, ..., T$ ($1 \leq T < \infty$). Each stage t consists of a simultaneous move of the players: each player chooses between 'cooperate' (C) and 'defect' (D). A stage-t outcome $(B^1, B^2) \in \{C, D\}^2$ leads

	C	D
C	2, 2	0, 3
D	3, 0	1, 1

	C	D
C	$2 + 2s, 2 + 2s'$	$3s, 3$
D	$3, 3s'$	$1 + s, 1 + s'$

Table 1: Monetary transfers (left) and utilities (right) at a stage in which player 1 (2) has state s (s')

to monetary transfers to the players of the structure of a prisoners' dilemma (Table 1, left). Let $v_{B^1 B^2}^i$ denote player i's transfer, or *material payoff*. For instance, $v_{CC}^1 = 2$. A player i's intra-period *utility* from (B^1, B^2) may also

be affected by the other player j's transfer, and this to an extent given by i's current sympathy level (state) $s \in \mathbf{S} = [0,1]$, where $s = 1$ represents full sympathy and $s = 0$ full self-interestedness.[10] Formally, let i's intra-period utility be $u^i(B^1, B^2; s) := v^i_{B^1 B^2} + s v^j_{B^1 B^2}$, the sum of i's own transfer and j's transfer weighted by current sympathy.[11] Player i's intertemporal utility of a path $h \equiv (B^1_t, B^2_t)_{t=0,\ldots,T} \in (\{C,D\}^2)^{T+1}$ is the sum of his intra-period utilities:

$$U^i(h) = \sum_{t=0}^{T} u^i(B^1_t, B^2_t; s^i(h_t)) = \sum_{t=0}^{T} \left[v^i_{B^1_t B^2_t} + s^i(h_t) v^j_{B^1_t B^2_t} \right], \tag{5}$$

where $s^i(h_t)$ denotes player i's sympathy state after experiencing the past moves $h_t := (B^1_t, B^2_t)_{t=0,\ldots,t-1}$.

Once again, the core question is: how should $s^i(h_t)$ be specified, i.e. how do experiences affect the sympathy state? Suppose the linear change model is used. The experiences of cooperation and defect by the other player are then identified with a strength-indexed sympathy state: $C \equiv s^C_{x^C}$ and $D \equiv s^D_{x^D}$. Initially, the players are each in the constitution $c \equiv s_x$, another strength-indexed state.[12] If the other player j cooperates at the initial stage, player i's sympathy state changes to $\overline{c|C} = \frac{xs + x^C s^C}{x + x^C}$. More generally, when reaching stage t, player i has experienced a sequence $B^j_0, \ldots, B^j_{t-1}$ of moves of the other player, leading to the new sympathy state

$$s^i(h_t) = \overline{c|B^j_0 \cdots |B^j_{t-1}} = \frac{xs + n_t x^C s^C + (t - n_t) x^D s^D}{x + n_t x^C + (t - n_t) x^D},$$

with n_t denoting the number of times cooperation C ($\equiv s^C_{x^C}$) occurs among $B^j_0, \ldots, B^j_{t-1}$. Note that we have now fully specified a dynamic game with endogenously changing mutual sympathy. For many reasonable parameter combinations[13], there exists a subgame perfect equilibrium such that (along the equilibrium path) both players cooperate at all stages. Interpretationally, a player cooperates in early stages[14] in order to win the sympathy of the other player (although cooperation gives him less *current* utility, and of course less monetary payoff), and later players cooperate because they like each other by then (with cooperation now being dominant in the current constituent game). This contrasts with the always-defect prediction in classical finitely repeated prisoners' dilemmas.[15]

---

[10] By letting $\mathbf{S} = [-1,1]$ one could also capture antipathy.

[11] Whenever I denote a player by i, I denote the other by j.

[12] Notice two implicit assumptions (that could be dropped). First, the parameters s^C, x^C, s^D, x^D, s, x are the same for each player. Second, a player's state is not affected by his *own* actions (which neglects phenomena such as habit-formation).

[13] Reasonably, $s^D < s < s^C$, i.e. a defect-experience reduces sympathy, and a cooperate-experience increases it. Also, the strengths of experience x^C, x^D and the number of periods T should not be too small, to leave sufficient potential for state change.

[14] The number of early stages is zero if initial sympathy c already exceeds $1/2$.

[15] It is worth mentioning *psychological (dynamic) games* as another fruitful approach for ex-

4 Two postulates about change and a consequence

In this and the next two sections, we consider a change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ for a *one-dimensional* state space $\mathbf{S}$. Although all notions and axioms introduced in these sections apply equally to multi-dimensional states, the reader is invited to restrict his or her attention to the one-dimensional case for now; multi-dimensional states are postponed to Section 8.

I now introduce two natural postulates on change – Attraction and Indoctrination – and prove that, on the background of a richness assumption, they imply a striking restriction: switching the order of two experiences has no effect on the state to which the individual is ultimately attracted. For instance, if as in Application 2 the individual is a player and his states are his sympathy levels for the opponent, then experiencing first cooperation and then defection by the opponent attracts the player to the same sympathy level as experiencing first defection and then cooperation. This conclusion is non-obvious because none of my postulates deals explicitly with the order of experience. The finding can be interpreted in two ways: either as a welcome argument for ignoring the order of experience when modelling change, a simplicity gain; or as a warning that modelling order-sensitive change behaviour requires giving up at least one of the basic assumptions.[16]

I start with the first postulate. Real-life experiences usually 'pull' us in some direction, 'suggest' to us to be of some kind: nice behaviour of Sam suggests liking Sam, drinking wine 'pulls' towards higher wine addiction, and so on. I formalise this using the notion of attraction to a state:

Definition 1 *An experience e* attracts to *a state s if for every constitution c the new state* $\overline{c|e}$ *is s or is strictly between s and the old state* $\bar{c}$. *An experience is* attracting *if it attracts to a state.*
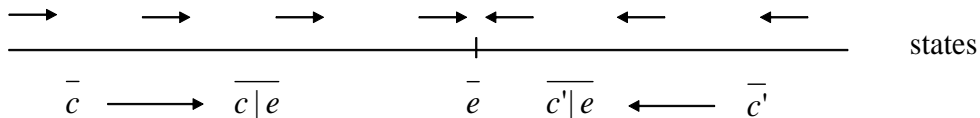

Figure 1: An experience e attracting constitutions c and c'

An attracting experience has the (plausible) feature of always moving the

---

plaining the emergence of cooperation and reciprocal behaviour (e.g. Geanakoplos et al. 1989). There, a player's sympathy levels (and hence, his utilities) can depend on his beliefs about other players' beliefs, and hence indirectly on his beliefs about the 'kindness' of their intentions.

[16]The co-existence of two readings – direct or by contraposition – pertains to many results, including Aumann's on agreeing to disagree. Aumann's (1976) celebrated result can be read either as supporting that agents do not 'agree to disagree', or as showing that modelling agents who 'agree to disagree' requires giving up a basic assumption (of common priors or of Bayesian updating).

agent's state towards the same point, regardless of where the agent started. Two facts are worth recording (the proofs are obvious).

- For any (initial) constitution c, if an experience e attracts to the old state $\bar{c}$, then the state does not change, i.e. $\overline{c|e} = \bar{c}$.
- Each experience e attracts to at most one state, which (if existent) is denoted $\bar{e}$ and called the *attractor* of e or simply the *state* of e.

The first postulate requires experiences to be of the attracting kind:

Attraction (A) Each experience e attracts to a state $\bar{e}$.

Attraction. which holds in our Examples 1 and 2[17], is a plausible, though not universal property of change. It notably allows an experience to attract to the maximal (resp. minimal) state in **S**, in which case the experiences *always* raises resp. reduces the individual's state.

The second postulate concerns the effect of repeated experience:

Indoctrination (I) For every experiences e, writing c_n for $c|e\cdots|e$ (the result of n times experiencing e),

(I_1) for any initial constitutions c, c', the difference in final state, $\overline{c_n} - \overline{c'_n}$, converges to zero as $n \to \infty$ (in short: unboundedly growing future experience ultimately overrules the past);

(I_2) for any initial constitution c, the effect of any experience e' on the final state, $\overline{c_n|e'} - \overline{c_n}$, converges to zero as $n \to \infty$ (in short: unboundedly growing past experience ultimately overrules the future).

Indoctrination (which again holds in Examples 1 and 2[18]), is another plausible but not universal property. Part of the plausibility lies in the fact that only an asymptotic requirement is made, and that asymptotic negligibility (of the past in I_1 and the future in I_2) is required only for the highly extreme and artificial circumstances in which the individual has (is 'indoctrinated' by) exactly the same experience e over and over again, without distraction by other experiences in between and without the 100th experiences being any different or weaker than the first. Intuitively, as total experience grows, the past (in I_1) or future (in I_2) matters less and less in comparison, and ultimately becomes negligible. This is plausible if we exclude *decay*: an experience e does not later gradually lose its power, it is not 'forgotten' as time progresses and further (possibly identical) experiences are made.[19]

---

[17] The attractor of an experience is given by $\overline{s_x} = s$ in Example 1 and by $\bar{e} = Median(e)$ in Example 2.

[18] For instance, I_1 holds in the linear model because, writing $c = s_x$ and $e = \tilde{s}_{\tilde{x}}$, one has $\overline{c_n} = \frac{xs + n\tilde{x}\tilde{s}}{x + n\tilde{x}} \to \tilde{s}$ and, similarly, $\overline{c'_n} \to \tilde{s}$. I_2 holds in Example 2 because $\overline{c_n} = Median(c + ne) \to Median(e)$ and $\overline{c_n|e'} = Median(c + ne + c') \to Median(e)$.

[19] Our framework is mainly aimed at change *without decay* of the effect of past experiences

Notice also that both previous conditions exclude the existence of a 'neutral' experience which leaves all states unchanged.

The composition of two experiences is naturally defined as follows.

Definition 2 *An experience $\hat{e}$ is the* composition *of experiences e, e' if $\hat{e}$ has the same effect as e followed by e', i.e. if $c|\hat{e} = c|e|e'$ for all constitutions c.*

An obvious remark follows from (1):

- For all experiences e, e', there is (i.e. **E** contains) *at most one* composition of e and e'; if there is one, it is denoted $e \circ e'$ or simply ee'.

My results will assume the set of experiences to be 'closed under composition':

Richness$_1$ (R$_1$) If **E** contains experiences e, e', it contains their composition ee'.

Richness$_1$ is satisfied in many examples.[20] A model violating R$_1$ can always be enriched to one satisfying R$_1$ by simply 'closing' **E** under composition.[21]

As shown in the appendix, composition of experience defines an *associative* operation on constitutions (given Richness$_1$), so that one may drop brackets without ambiguity: $ee'e''$ stands for either $e(e'e'')$ or $(ee')e''$, and e^n for the n-fold self-composition $e \cdots e$ $(n \geq 1)$.

Some brief remarks about the role of richness conditions in axiomatics are due. Virtually all formal models in decision theory have their own richness conditions; e.g. Savage's and von-Neumann-Morgenstern's models assume the agent to face a rich set of acts resp. lotteries.[22] This paper uses certain conditions of richness in experiences or constitutions. If in a concrete application the agent cannot 'have' all these experiences or constitutions (because they simply do not occur, are 'infeasible' in the special environment), then our rich model refers to an extension

---

and initial states and constitutions. Accordingly, in Indoctrination, each new occurrence of e is intuitively 'added' to the stock of earlier ones, without 'replacing' or 'diminishing' them. Extending our approach so as to allow for decay is a challenge left for future research and might be accomplished in various ways, some of which involve weakening or dropping Indoctrination. One way (which is compatible with retaining Indoctrination) explicitly augments change models by *decay* (or '*de-experience*') operators, which transform individual constitutions in the opposed direction from experiences: they 'undo' the effect of experience. In group-theoretic terms, they are *inverses* of experiences relative to composition (see Definition 2). Change models as currently defined do not allow inversion of experiences: $(\mathbf{E}, \circ)$ is just a *semi*group, as proven later.

[20] In our Examples 1 and 2, composition is given by $s_x \tilde{s}_{\tilde{x}} = \left(\frac{sx + \tilde{s}\tilde{x}}{x + \tilde{x}}\right)_{x + \tilde{x}}$ and $e\tilde{e} = e + \tilde{e}$, respectively.

[21] W.l.o.g., identify experiences in **E** with transformations $\mathbf{C} \to \mathbf{C}$. Extend **E** to a set $\hat{\mathbf{E}}$ by adding all those transformations $\mathbf{C} \to \mathbf{C}$ that are compositions of (two or more) transformations in **E**, and extend the revision rule $\mathbf{C} \times \mathbf{E} \to \mathbf{C}$ to a revision rule $\mathbf{C} \times \hat{\mathbf{E}} \to \mathbf{C}$ in the obvious way.

[22] The set of Savage acts (mappings from nature states to outcomes) is closed under mixing and contain all 'constant acts'. Von-Neumann-Morgenstern's agent chooses from the set of *all* lotteries (over given deterministic outcomes).

of the real environment so as to also include what *would* happen in hypothetical cases.[23]

Theorem 1 *If a change model* $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ *for a one-dimensional state space* $\mathbf{S}$ *satisfies Attraction, Indoctrination and Richness$_1$, then* $\overline{ee'} = \overline{e'e}$ *for all experiences* $e, e' \in \mathbf{E}$.

This order-invariance finding can be illustrated by the models of Examples 1 and 2. In Example 1, $\overline{s_x \tilde{s}_{\tilde{x}}} = \overline{\tilde{s}_{\tilde{x}} s_x} = \frac{xs + \tilde{s}\tilde{x}}{x + \tilde{x}}$ for all experiences $s_x, \tilde{s}_{\tilde{x}}$; and in Example 2, $\overline{ee'} = \overline{e'e} = Median(e + e')$ for all experiences e, e'. As Theorem 1 shows, the order-invariance property is no coincidence: it holds not just for change models as simple as Examples 1 and 2, but for *all* change models satisfying Attraction, Indoctrination and Richness$_1$. Order-invariance in the full sense would amount to $ee' = e'e$ rather than just $\overline{ee'} = \overline{e'e}$ for all experiences e, e'. Full order-invariance says not only that the two experiences ee' and $e'e$ attract to the same attractor state (perhaps differently strongly), but that they are *the same* – a far stronger conclusion which also implies that they attract the initial constitution equally strongly and indeed that the new constitution is the same both times, so that any next experience e'' will have the same effect both times, and so on. Full order-invariance is obtained in Theorem 2 below.

To illustrate the proof given in the appendix, consider any experiences e, e'. The proof brings to light two intuitive facts about any experience g and its n-fold repetition $g^n = g \cdots g$, where $n \geq 1$. Firstly, g^n attracts to the same state as g. Formally, $\overline{g^n} = \overline{g}$ for all $n \geq 1$. Secondly, in the composite experience $eg^n e'$ the effect of the parts e and e' vanishes compared to the effect of the part g^n as n gets very large, so that asymptotically the experience attracts to the same state as g^n, i.e. as g. Formally, $\overline{eg^n e'} \to \overline{g}$ as $n \to \infty$. Applying the second mentioned fact with $g := e'e$, we obtain $\overline{e(e'e)^n e'} \to \overline{e'e}$ as $n \to \infty$. Now $\overline{e(e'e)^n e'}$ is the same as $\overline{(ee')^{n+1}}$ by associativity of composition, and hence as $\overline{ee'}$ by the first mentioned fact (applied with $g := ee'$). Thus, $\overline{ee'} \to \overline{e'e}$ as $n \to \infty$, so that (since $\overline{ee'}$ does not depend on n) $\overline{ee'} = \overline{e'e}$, as desired.

5   Strength of constitution and strength of experience

The above postulates – Attraction and Indoctrination – might be viewed as defining the large class of 'prima facie plausible' change models, which includes models as different as the linear model and Example 2's non-parametric model. Within this class, the linear model deserves our special attention: it is probably the simplest (interesting) model, and it has something very compelling to it in that the

---

[23] A less rich model $(\hat{\mathbf{S}}, \hat{\mathbf{E}}, \hat{\mathbf{C}}, (\hat{\bar{\cdot}}), (\hat{.|.}))$ automatically inherits all findings about a rich extension $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$, such as the commutativity of experience. (Model extensions are defined in Section 2.) So the paper's findings also have a bearing on less rich environments.

individual's post-experience state is a weighted average of where he was before and where the experience wants him to be. But what exactly (if anything) makes the linear model so special among the class of 'prima facie plausible' change models? It is a single additional property, *Attraction-Consistency*, as proven in the next section. In the present section, I introduce Attraction-Consistency and prove two consequences of this condition (in conjunction with the previous postulates), namely in Theorem 2 that experience is fully commutative, and in Theorem 3 that, in short, the modeller is allowed to represent experiences and also constitutions as state-strength pairs $s_x \equiv (s, x)$, just as done in the linear model. Theorems 2 and 3 can again be interpreted in either normative or purely logical terms.[24]

The just-announced third condition on change states as follows.

Attraction-Consistency (AC). This condition has two parts.
($\mathbf{AC_1}$) For all experiences e, e' attracting to a same state s, if *some* constitution c is more attracted by e than by e' (i.e. the state $\overline{c|e}$ is strictly between s and $\overline{c|e'}$ or $\overline{c|e} = s \neq \overline{c|e'}$) then *each* constitution c with $s \neq \overline{c|e'}$ is so.
($\mathbf{AC_2}$) For all constitutions c, c' with the same state, if *some* experience e attracting to a state s attracts c more than c' (i.e. the state $\overline{c|e}$ is strictly between s and $\overline{c'|e}$ or $\overline{c|e} = s \neq \overline{c'|e}$) then *each* experience e attracting to a state $s \neq \overline{c'|e}$ does so.

Condition AC is, more than A and I, a genuine restriction of generality (and a cornerstone on the way towards linearity of chance, as will turn out later). While satisfied by the linear model, AC fails for Example 2's non-parametric model. AC_1 states that any two experiences e, e' which attract to the same state can be unambiguously compared in terms of their strength, in the sense that if e is stronger than e' 'sometimes' (i.e., in the effect on 'some' constitution), then e is stronger 'always'. If for instance states are levels of altruism and e and e' both attract towards a given state of high altruism, then it cannot be that e raises altruism more strongly than e' whenever the agent starts at low altruism, while e' raises altruism more strongly than e whenever the agent starts at medium altruism. Similarly, AC_2 states that any two constitutions c, c' with same state can be unambiguously compared in terms of their strength, in the sense that if c' is stronger than c 'sometimes' (i.e., resists better to 'some' experience), then c' is stronger 'always'.

The section's theorems require a second richness condition. I call a constitution c_w *weak* if every attracting experience e *fully* attracts c_w, i.e. $\overline{c_w|e} = \overline{e}$. Intuitively, the agent in a weak constitution does not resist at all to any experience, obviously

---

[24]The theorems can be interpreted *either* as providing arguments for treating experience as commutative and modelling experiences and constitutions as state-strength pairs, *or* as informing us that we cannot violate these features unless we sacrifice Attraction or Indoctrination or Attraction-Consistency.

a limiting type of constitution. In Example 1, the weak constitutions are the zero-strength constitutions s_0 ($s \in \mathbf{S}$). In Example 2, the only weak constitution is the zero-measure on $\mathbf{S}$.

Richness$_2$ (R$_2$) For every non-weak constitution c there is a weak constitution c_w from which c is reachable, i.e. such that $c = c_w|e_1 \cdots |e_n$ for some experiences $e_1, ..., e_n$ ($n \geq 1$).

This condition (which holds in Examples 1 and 2) is fairly plausible. Indeed, if someone starts in a weak constitution (provided there is at least one) it should intuitively be possible for him to reach any non-weak constitution through appropriate experiences – because weakness of constitution stands for the absence of any predispositions whatsoever, hence for the ability to be entirely shaped by experience. (This intuition is underscored by later lemmas.)

The conjunction of R$_1$ and R$_2$ is called Richness$_{1,2}$ (in symbols: R$_{1,2}$), and later notation should be interpreted similarly (e.g. Richness$_{1\text{-}3}$ stands for the conjunction of three richness conditions).

Theorem 2 *If a change model* $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ *for a one-dimensional state space* $\mathbf{S}$ *satisfies Attraction, Indoctrination, Attraction-Consistency and Richness$_{1,2}$, then* $ee' = e'e$ *for all experiences* $e, e' \in \mathbf{E}$.

Note the progress over Theorem 1: while Theorem 1 merely obtained that the order of experience is irrelevant for the ultimate point of attraction ('$\overline{ee'} = \overline{e'e}$'), Theorem 2 obtains full order-irrelevance ('$ee' = e'e$'), so that, intuitively speaking, the order is also irrelevant for the strength of this attraction. In algebraic terminology, $(\mathbf{E}, \circ)$ is an Abelian (i.e. commutative) semigroup. The proof is quite technical and does not lend itself to a short description. Nonetheless, the following observation might give the reader an intuition. What does it mean for the compound experiences ee' and $e'e$ to be *identical*. Intuitively, being identical amounts to two properties: firstly, the two experiences pull the agent towards the same state (i.e., $\overline{ee'} = \overline{e'e}$), and secondly, they do so equally strongly. The first of these properties is already obtained in Theorem 1, without yet assuming AC. The second property draws on the notion of strength of experience, a notion which is made unambiguous by introducing AC (whereas without AC the notion can be ambiguous; see fn. 2).

To state the next theorem (on the structure of experiences and constitutions), I now formally define strength comparisons:

Definition 3 *For every state s, let $\mathbf{C}_s$ be the set of constitutions with state s and $\mathbf{E}_s$ the set of experiences attracting to state s, and define the*

(a) strength relation $\geq$ *on* $\mathbf{E}_s$ *by:* $e \geq e'$ *("e is at least as strong as e'") if e attracts constitutions as least as much as e', i.e. for every constitution c, $\overline{c|e}$ is weakly between s and $\overline{c|e'}$;*

16

(b) strength relation $\geq$ *on* $\mathbf{C}_s$ *by:* $c \geq c'$ *("c is* at least as strong *as* c'*") if c is at most as attracted by experiences as* c'*, i.e. for every attracting experience* e, $\overline{c'|e}$ *is weakly between* $\overline{e}$ *and* $\overline{c|e}$.

The linear model, for instance, has $\mathbf{E}_s = \{s_x : x > 0\}$ and $\mathbf{C}_s = \{s_x : x \geq 0\}$, with strength relation on $\mathbf{E}_s$ (resp. $\mathbf{C}_s$) simply given by $s_x \geq s_{\tilde{x}} \Leftrightarrow x \geq \tilde{x}$, which is in line with our earlier interpretation of the 'x' in s_x as measuring strength of experience (resp. constitution).

The strength relation $\geq$ (on $\mathbf{E}_s$ resp. $\mathbf{C}_s$) induces a '(strictly) stronger than' relation $>$ and an 'as strong as' relation $\equiv$ (on $\mathbf{E}_s$ resp. $\mathbf{C}_s$), both defined as usual.[25] Endowing $\mathbf{E}_s$ and $\mathbf{C}_s$ with their strength relations yields ordinal structures $(\mathbf{E}_s, \geq)$ and $(\mathbf{C}_s, \geq)$. Further endowing $\mathbf{E}_s$ with composition $\circ$ (under R_1) yields a structure $(\mathbf{E}_s, \geq, \circ)$ (a so-called ordered semi-group, as we will see). Isomorphisms between structures (i.e. between sets endowed with relation(s) and/or operation(s)) are defined as usual, namely as relation- and operation-preserving bijections. Two structures $(A, ...)$ and $(B, ...)$ are *isomorphic* (written $(A, ...) \equiv (B, ...)$) if there exists an isomorphism between them. Isomorphic structures are thus identical up to relabelling.

We are now ready for the section's second result. I call a change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ *trivial* if its revision rule is constant, i.e. if the individual is changed to the same constitution $c^* = c|e$ whatever the initial constitution c and experience e. It follows by (1) that there is at most one experience, i.e. that $\#\mathbf{E} \leq 1$, and by (2) that constitutions are isomorphic to states, i.e. that one may assume w.l.o.g. that $\mathbf{C} = \mathbf{S}$ and $\overline{c} = c$ for all $c \in \mathbf{C}$.

Theorem 3 *If a non-trivial change model* $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ *for a one-dimensional state space* $\mathbf{S}$ *satisfies Attraction, Indoctrination, Attraction-Consistency and Richness$_{1,2}$, then for every state* $s \in \mathbf{S}$ *there exists a set* $X_s \subseteq (0, \infty)$ *(of 'strength levels') closed under addition such that*

- $(\mathbf{E}_s, \geq, \circ)$ *is isomorphic to* $(X_s, \geq, +)$, *and*
- $(\mathbf{C}_s, \geq)$ *is isomorphic to* $(X_s, \geq)$ *(hence, to* $(\mathbf{E}_s, \geq)$*) if* $\mathbf{C}_s$ *contains no weak constitution, and to* $(X_s \cup \{0\}, \geq)$ *if* $\mathbf{C}_s$ *contains a weak constitution.*

The set of strength levels X_s (for a state s) might for instance be $(0, \infty)$ (as in the linear model) or $[1, \infty)$ or $(0, \infty) \cap \mathbb{Q}$ or $\{1, 2, ...\}$ or $\{mx+ny : m, n \in \{1, 2, ...\}\}$ (for fixed $x, y > 0$). In fact, for every non-empty set $X \subseteq (0, \infty)$ closed under addition, a submodel of the linear model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ is obtained by replacing $\mathbf{E}$ and $\mathbf{C}$ by their subsets $\mathbf{S} \times X$ resp. $\mathbf{S} \times (X \cup \{0\})$ and restricting $(\bar{\cdot})$ and $(.|.)$ accordingly. Applying Theorem 3 to this submodel, the set X_s can be identified with X for all $s \in \mathbf{S}$. In other examples, the set X_s varies across states s.

---

[25] These relations on $\mathbf{E}_s$ are given by $e > e' \Leftrightarrow [e \geq e' \& e' \not\geq e]$ and $e \equiv e' \Leftrightarrow [e \geq e' \& e' \geq e]$; and similarly for the relations on $\mathbf{C}_s$.

Under Theorem 3's conditions, it is justified to represent experiences and also constitutions as state-strength pairs $s_x \equiv (s, x)$, with state projection given by $\overline{s_x} = s$, strength comparisons (between experiences or between constitutions) given by $s_x \geq s_{x'} \Leftrightarrow x \geq x'$, and composition of experiences given by $s_x s_{x'} = s_{x+x'}$. This brings us partially to the linear model. Theorem 4 will bring us there fully, by forcing all sets X_s to be $(0, \infty)$ and the revision rule to be linear.

The proof of Theorem 3 in the appendix begins by establishing that the algebraic structure $(\mathbf{E}_s, \geq, \circ)$ (for $s \in \mathbf{S}$) is an ordered semigroup which satisfies several salient properties, such as the properties of being Archimedean, commutative and cancellative. Applying Hölder's (1901) seminal representation theorem, we can then embed $(\mathbf{E}_s, \geq, \circ)$ into the ordered semigroup of positive reals, $((0, \infty), \geq, +)$. This proves part (a), and after some arguments also part (b).

6   Characterisation of the linear change model

As mentioned, the linear model deserves our special attention as it is the perhaps simplest and intuitively most natural (non-degenerate) change model. Does it have a compelling characterisation in terms of few easily interpretable properties? I now show that the linear model is, up to isomorphism, the *only* change model that satisfies our earlier conditions and is 'sufficiently rich' in experiences and constitutions (in the sense of five richness conditions). Formally, a change model $(\hat{\mathbf{S}}, \hat{\mathbf{E}}, \hat{\mathbf{C}}, (\hat{\cdot}), (.\hat{|}.))$ is *isomorphic to* (or a *reparametrisation of*) another one $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ if there exist an increasing bijection between states $\mathbf{S} \to \hat{\mathbf{S}}, s \mapsto s^*$, a bijection between constitutions $\mathbf{C} \to \hat{\mathbf{C}}, c \mapsto c^*$, and a bijection between experiences $\mathbf{E} \to \hat{\mathbf{E}}, e \mapsto e^*$, such that $(\hat{\cdot})$ is the image of $(\bar{\cdot})$ (i.e. $\overline{c}^* = \widehat{c^*}$ for all $c \in \mathbf{C}$) and $(.\hat{|}.)$ is the image of $(.|.)$ (i.e. $(c|e)^* = c^* \hat{|} e^*$ for all $c \in \mathbf{C}$ and all $e \in \mathbf{E}$).[26] Isomorphic models are perfectly equivalent (but perhaps not equally natural or convenient).[27]

Here are the first two additional richness conditions characteristic of linear models:

Richness₃ (R₃) For each constitution c, some experience e leaves the state unchanged, i.e. $\overline{c|e} = \overline{c}$.

Richness₄ (R₄) No non-weak constitution c is weaker than all other non-weak constitutions with the same state as c, i.e. satisfies $c < c'$ for all other non-weak constitutions c' with state $\overline{c'} = \overline{c}$.

---

[26] For instance, the linear model for the state space $\mathbf{S} = [0, 1]$ is isomorphic to that for the state space $\hat{\mathbf{S}} = [0, 2]$: transform states via $s \mapsto 2s$, constitutions via $s_x \mapsto (2s)_x$ and experiences via $s_x \mapsto (2s)_x$. Linear models can also be reparameterised into non-linear models, e.g. by measuring strength of experience on a new, logarithmic scale.

[27] From a formal angle, 'is isomorphic to' defines an equivalence relation over the class of change models.

Intuitively, R_3 requires the set of experiences to be rich enough that any state is confirmed by at least one experience.[28] To illustrate R_4 using the linear model, note that for each non-weak constitution s_x $(x \neq 0)$ we may consider the weaker constitution $s_{\frac{1}{2}x}$. In general, R_4 requires the set of constitutions to be sufficiently rich that for any non-weak constitution c there is another non-weak constitution with the same state which is not stronger than c, i.e., either weaker than c or as strong as c or non-comparable with c in terms of the strength relation. If the strength relation over all constitutions with the given state is a linear order (which it is under A, I, AC and $R_{1,2}$ by Lemma 11), then the last two possibilities of equal strength and non-comparability disappear, so that R_4 can be characterized more straightforwardly as requiring that for any non-weak constitution c there is a weaker non-weak constitution. So, if for instance states are altruism levels, then for any (non-weak) constitution with low altruism the agent can be in a (non-weak) constitution with the same low altruism but less resistance to experience.

To state the last richness condition, I define a *state path* as a family $(s_{\mathbf{e}})_{\mathbf{e} \in \cup_{n=0}^{\infty} \mathbf{E}^n}$ $(\in \mathbf{S}^{\cup_{n=0}^{\infty} \mathbf{E}^n})$ of states $s_{\mathbf{e}} \in \mathbf{S}$ assigned to experience sequences $\mathbf{e} \equiv (e_1, ..., e_n) \in \mathbf{E}^n$ of any (possibly zero) length n. A state path $(s_{\mathbf{e}})_{\mathbf{e} \in \cup_{n=0}^{\infty} \mathbf{E}^n}$ describes where the individual is initially (namely in state $s_{()}$), after any experience e (namely in state $s_{(e)}$), after any pair of experiences e_1, e_2 (namely in state $s_{(e_1, e_2)}$), and so on. To each constitution c is naturally assigned state path $(\overline{c|e_1 \cdots |e_n})_{(e_1,...,e_n) \in \cup_{n=0}^{\infty} \mathbf{E}^n}$, containing the initial state $\overline{c}$, the states $\overline{c|e}$ after any experiences e, and so on.[29] Of course, a state path in $\mathbf{S}^{\cup_{n=0}^{\infty} \mathbf{E}^n}$ need not be possible, i.e. need not pertain to any constitution in $\mathbf{C}$. A state path is *constant* if its states are all the same, i.e. if the agent 'never changes'.

Richness$_5$ (R$_5$) For any sequence of constitutions $(c_k)_{k=1,2,...}$, if the sequence of corresponding state paths converges (pointwise) to a non-constant state path, then there is a constitution c with this state path.

Intuitively, R_5 requires $\mathbf{C}$ to be closed under taking 'limiting constitutions'. In R_5, the constitution c is indeed the limit of the sequence $(c_k)_{k=1,2,...}$ in the sense of a natural topology.[30] Another perspective on R_5 is that it requires topological closedness (in fact, slightly less than closedness due to the qualification 'non-constant') of the set of constitutions $\mathbf{C}$ as embedded into the state path space

---

[28]Theorem 4 still holds if we weaken R_3 by restricting it to ('extreme') constitutions c whose state $\overline{c}$ belongs to the boundary of $\mathbf{S}$ – a mild richness condition, it seems.

[29]In fact, a constitution is fully characterized by its state path (by (1)). Formally speaking, we could thus identify constitutions with their state paths.

[30]Here, I endow $\mathbf{C}$ with the *weak topology* induced by the functions $f_{\mathbf{e}} : \mathbf{C} \to \mathbf{S}$ (with $\mathbf{e} \equiv (e_1, ..., e_n)$ ranging over $\cup_{n=0}^{\infty} \mathbf{E}^n$) defined by $f_{\mathbf{e}}(c) = \overline{c|e_1 \cdots |e_n}$. This topology is by definition the smallest (coarsest) topology for which these functions are continuous. So, a constitution sequence (c_n) converges to a constitution c if and only if $f_{\mathbf{e}}(c_n) \to f_{\mathbf{e}}(c)$ for all functions $f_{\mathbf{e}}$, or equivalently, if and only if c_n's state path converges (pointwise) to c's state path as $n \to \infty$.

$\mathbf{S}^{\cup_{n=0}^{\infty} \mathbf{E}^n}$.[31]

We are ready for the characterization result.

Theorem 4 *A change model* $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\cdot), (.|.))$ *for a one-dimensional state space* $\mathbf{S}$ *is the linear model (up to isomorphism) if and only if it satisfies Attraction, Indoctrination, Attraction-Consistency and Richness$_{1-5}$.*

The proof of this theorem in the appendix again begins by analysing the structure $(\mathbf{E}_s, \geq, \circ)$, i.e. the set of experiences attracting to a given state s endowed with the strength relation and the composition operation. Using Theorem 4's assumptions, this structure turns out to be an ordered semigroup with the properties of density and completeness and the property of continuity of the operation $\circ$, where these three properties are understood in the topological (rather than ordinal) sense and refer to the so-called order topology induced by $\geq$. We can then apply an important representation theorem by Arzél (1948) and Tamari (1949), which, after additional arguments, tells us that $(\mathbf{E}_s, \geq, \circ)$ is isomorphic to the ordered semi-group of positive reals, $((0, \infty), \geq, +)$ (rather than being merely *embeddable* into $((0, \infty), \geq, +)$, the property obtained in Theorem 3 under weaker richness assumptions). Similarly, the structure $(\mathbf{C}_s, \geq)$, i.e. the set of constitutions with state s endowed with the strength relation, turns out to be isomorphic to $([0, \infty), \geq)$. These findings take us 'half way' towards a linear change model. The other half of the proof analyses (among other things) composition as an operation on $\mathbf{E}$ (rather than on some $\mathbf{E}_s$); this operation is shown to be isomorphic to composition in the linear model. More details of the proof strategy are given in the appendix.

7 Two further applications

The following two applications serve a twofold goal. First, they illustrate how change models can help to capture changes in fundamental preferences ('dynamic inconsistency'), as opposed to the less radical kinds of changes already illustrated in Applications 1 and 2 above. Second, they are typical examples in which the characteristic in question can easily be multi-dimensional – a scenario formally analysed in the next section.

Application 3: asymmetric information about instable players. There are many interesting dynamic games in which some players – call them *instable* players – have a characteristic that (i) changes in the course of the game and (ii) is

---

[31]This embedding relies on identifying constitutions with their state paths (a one-to-one mapping by (2)). 'Closedness' is meant relative to the pointwise-convergence topology on $\mathbf{S}^{\cup_{n=0}^{\infty} \mathbf{E}^n}$ (the weak topology induced by the projection functions, i.e. by the functions $\mathbf{S}^{\cup_{n=0}^{\infty} \mathbf{E}^n} \to \mathbf{S}$ that evaluate state paths at particular points $\mathbf{e} \in \cup_{n=0}^{\infty} \mathbf{E}^n$).

preference-relevant in that the states of the characteristic at some point(s) of time affect the utility of the player or perhaps of another player. Players with changing sympathy levels (Application 2) are just one example. Another example are dynamically inconsistent players whose preference over outcomes depends on the decision node. It is often realistic to assume that there is incomplete information about

   (i) an instable player's initial constitution (how much wine-addiction does he initially have? how much sympathy?), and/or
  (ii) the effect of the player's moves on his constitution (how does his wine-drinking affect his wine-addiction?) or the effect of other players' or nature's moves on his constitution (how does cooperation of other players affect his sympathy level?).

More formally, using the linear change model (in which constitutions and experiences are represented by strength-indexed states), there may be incomplete information about (i) an instable player's initial constitution $c \equiv s_x$ and/or (ii) the precise experiences $s^A_{x^A}$ certain moves A in the game give him. Note that not just the other players may be incompletely informed about a given instable player, but also (or only) the player himself: sometimes we are the worst judges of ourselves. The relevance of such games is obvious.

Application 4: explaining dynamic inconsistency by change in characteristics. As mentioned, dynamic inconsistency is change in *fundamental* preference, i.e. preference over maximally described outcomes (as opposed to Applications 1 and 2, to Becker's theory, and to information-driven preference change[32]). Models of dynamically inconsistent agents often suffer from empirical underdetermination and an abundance of free parameters. In response, a change model could be used to constrain ('discipline') preference change. To see how, denote by $\mathcal{A}$ the set of relevant alternatives (e.g. terminal histories of a decision tree or dynamic game form) and represent the individual as holding at any moment (e.g. any decision node) a preference relation $\succeq_s$ on $\mathcal{A}$ that is fully determined by the current state $s \in \mathbf{S}$ of some given characteristic (such as drug addiction, criminal energy, health, or altruism). This explains change in preference by change in that characteristic, which (using a change model) is in turn explained by experiences $e \in \mathbf{E}$ such as drug consumption, (un)friendly actions of others, or internal experiences like Alzheimer or puberty. As in Applications 1-3, the question is once again: which change model should be used?

8 Change in multi-dimensional characteristics

While our previous axiomatic treatment (Sections 4-6) has focused on one-dimensional states, this section turns to a change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ for a Euclidean state

---

[32]I.e. belief-driven change in expected utilities (of non-fully described outcomes).

space $\mathbf{S} \subseteq \mathbb{R}^n$ of an arbitrary dimensionality $n \geq 1$. The dimensions could for instance represent taste for different goods. As it turns out, the entire previous analysis – the definitions, axioms and theorems – extend without modifications or complications, once we appropriately clarify the meaning that our definition of 'attracting' takes on in the now n-dimensional context. Recall that an experience e is said to *attract* towards a state s ($= \bar{e}$) if for each initial constitution c the new state $\overline{c|e}$ is s or is strictly between s and the old state $\bar{c}$. What means being (weakly or strictly) 'between' two states s and s' in the n-dimensional case? It naturally means belonging to the straight line joining s and s' (with or without the endpoints, respectively).[33] With this clarification in mind, all axioms are well-defined, and literal n-dimensional generalizations of Theorems 1-4 hold, as stated in a moment. From an interpretational perspective, however, the requirement that experiences are 'attracting' (i.e., the axiom Attraction) seems far more demanding in the multi-dimensional case $n > 1$ than in the one-dimensional case. For an experience e to attract a constitution c to the state $\bar{e}$, the state must move towards $\bar{e}$ along the line segment joining the old state $\bar{c}$ and $\bar{e}$; so, the state is attracted *proportionally in each dimension* (Figure 2, left).But many real-life ex-

An experience e attracting to $\bar{e}$ An experience e attracting only in one dimension
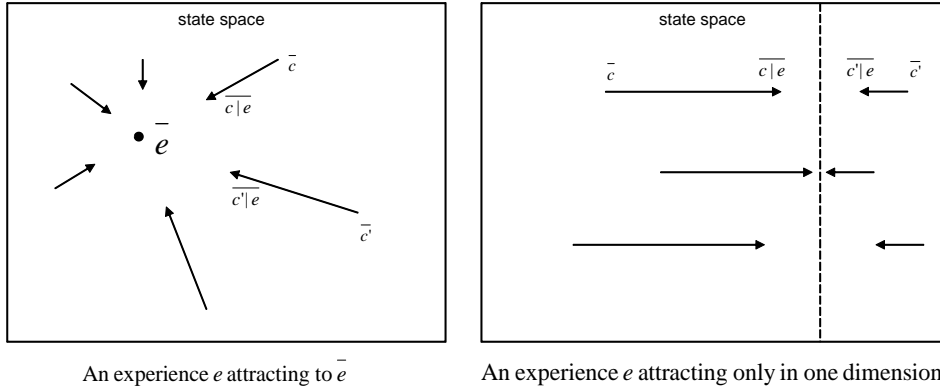
Figure 2: Our strong notion of 'attracting' (left) and a weaker notion (right) for a two-dimensional state space

periences attract more strongly in one dimension than in another, or even only in one dimension (Figure 1, right). If the first dimension represents taste for Bach music and the second taste for bananas (where $n = 2$), then the experience of a Bach concert might affect the state only in the first dimension. Our first axiom (Attraction) rules out such experiences.

The n-dimensional generalizations of Theorems 1-4 state as follows:

Theorem 1* *If a change model* $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ *for an n-dimensional state space* $\mathbf{S}$ *($n \geq 1$) satisfies Attraction, Indoctrination and Richness$_1$, then* $\overline{ee'} = \overline{e'e}$

_____

[33]The straight line with endpoints is the set $[s, s'] := \{\lambda s + (1-\lambda)s' : 0 \leq \lambda \leq 1\}$. The straight line without endpoints is the set $[s, s']\backslash\{s, s'\}$.

for all experiences $e, e' \in \mathbf{E}$.

Theorem 2* *If a change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ for an n-dimensional state space $\mathbf{S}$ ($n \geq 1$) satisfies Attraction, Indoctrination, Attraction-Consistency and Richness$_{1,2}$, then $ee' = e'e$ for all experiences $e, e' \in \mathbf{E}$.*

Theorem 3* *If a non-trivial change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ for an n-dimensional state space $\mathbf{S}$ ($n \geq 1$) satisfies Attraction, Indoctrination, Attraction-Consistency and Richness$_{1,2}$, then for every state $s \in \mathbf{S}$ there exists a set $X_s \subseteq (0, \infty)$ (of 'strength levels') closed under addition such that*

- *$(\mathbf{E}_s, \geq, \circ)$ is isomorphic to $(X_s, \geq, +)$, and*
- *$(\mathbf{C}_s, \geq)$ is isomorphic to $(X_s, \geq)$ (hence, to $(\mathbf{E}_s, \geq)$) if $\mathbf{C}_s$ contains no weak constitution, and to $(X_s \cup \{0\}, \geq)$ if $\mathbf{C}_s$ contains a weak constitution.*

Theorem 4* *A change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ for an n-dimensional state space $\mathbf{S}$ ($n \geq 1$) is the linear model (up to isomorphism) if and only if it satisfies Attraction, Indoctrination, Attraction-Consistency and Richness$_{1\text{-}5}$.*

As already indicated, if $n > 1$ then this axiomatic analysis of change is limited to rather special experiences, due to the demanding nature of the axiom of Attraction if $n > 1$. A more general treatment of multi-dimensional change goes beyond the scope of this paper. It would presumably have to weaken the notion of 'attracting' by allowing for different degrees of attraction along different dimensions. One might conjecture that, once our axiomatic analysis is appropriately generalized and based on a weaker notion of 'attracting', analogues of our four theorems can be obtained.[34]

Interestingly, if $n > 1$ then in Theorem 4* the isomorphism to the linear model holds in a particularly strong sense: the change model can be 'linearized' without transforming the state.[35] This additional structural property distinguishes the multi-dimensional case $n > 1$ from the one-dimensional case $n = 1$; it arises as a complex consequence of the 'interplay of dimensions'.

Theorems 1*-4* can be derived as corollaries of Theorems 1-4. The key insight needed to apply the 'one-dimensional' theorems to the n-dimensional case is as follows. Consider an n-dimensional state space $\mathbf{S}$ and a change model $\mathcal{M} := (\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ satisfying Attraction. $\mathcal{M}$ induces a sub-model $\mathcal{M}_T$ for every convex (sub-) state space $T \subseteq \mathbf{S}$. This sub-model, denoted $\mathcal{M}_T$, is defined by

---

[34]I conjecture that the order-invariance results of the first two theorems are still obtainable. As for the third theorem, I conjecture that the 'strength levels' of experiences or constitutions have to be defined as n-dimensional vectors rather than single numbers, since an experience or constitution can be differently strong in different dimensions. Similarly, I conjecture that the fourth theorem has an analogue involving a generalization of the linear model in which strength of experience or constitution is now n-dimensional.

[35]Formally, the state transformation $s \mapsto \widehat{s}$ can be chosen to be the identity function.

restricting the original model to T: the state space is of course T, the set of experiences is $\mathbf{E}_T = \{e \in \mathbf{E} : \bar{e} \in T\}$, the set of constitutions is $\mathbf{C}_T = \{c \in \mathbf{C} : \bar{c} \in T\}$, the state projection is that of $\mathcal{M}$ now restricted to $\mathbf{C}_T$, and the revision rule is that of $\mathcal{M}$ now restricted to $\mathbf{C}_T \times \mathbf{E}_T$. If T is chosen to be a line segment $T = [a, b] = \{\lambda a + (1 - \lambda)b : \lambda \in [0, 1]\}$ (for distinct states $a, b \in \mathbf{S}$), then it is *essentially one-dimensional*.[36] So the sub-model $\mathcal{M}_{[a,b]}$ falls into the scope of Theorems 1-4, since these theorems of course also hold for essentially rather than properly one-dimensional state spaces. One can prove Theorems 1*-4* by applying Theorems 1-4 'locally', i.e. to appropriately chosen sub-models $\mathcal{M}_{[a,b]}$ ($a, b \in \mathbf{S}$, $a \neq b$). More details are given in the appendix.

9 Conclusion

I have developed a systematic approach to modelling change in an agent's characteristics, as a step towards filling the wide gap between recognised importance of change and lack of theoretical understanding of it. The findings can be applied in many ways:

- The decision- or game-theorist might model either 'orthodox' change which keeps fundamental preferences fixed (see Applications 1, 2 and 3), or 'unorthodox' change which induces dynamic inconsistency (see Applications 3 and 4).
- He might either take our theorems as reasons for neglecting the order of experience (Theorems 1 and 2) and perhaps modelling change linearly (Theorems 3 and 4), or he might insist on order-relevance and non-linearity, which forces him to abandon Attraction, Indoctrination or Attraction-Consistency.
- The empirical researcher might estimate the real-life value of parameters of a given change model, such as the strength of the experience of cooperation by other people (players).

There is plenty of room for follow-up work: one could study other conditions on change models, generalize our initial treatment of multi-dimensional states (in order to better understand simultaneous change in interrelated characteristics like feelings for one's partner and pleasure at work), introduce the possibility of decay in the long-term effect of experiences and initial constitutions, study various dynamic games with change in individual characteristics, and so on.

10 References

AUMANN, R. (1976) Agreeing to disagree, *Ann. Statist.* 4(6): 1236-1239

ALIMOV, N. G. (1950) On ordered semigroups, *Izv. Akad. Nauk SSSR. Ser.*

---

[36] In the sense that a simple linear re-parametrisation transforms T into a real interval $T' \subseteq \mathbb{R}$; for instance, transform each state $s = \lambda a + (1 - \lambda)b \in T$ into $\lambda \in T' = [0, 1]$.

Mat. 14: 569-576 (Russian); *Math. Rev.* 12: 480 (English translation by R. P. Rich, Johns Hopkins University)

ARZEL, J. (1948) Sur les opérations définies pour nombres réels, *Bull. Soc. Math. France* 76: 59-64

BECKER, G. S. (1996) *Accounting For Tastes*, Harward University Press, Cambridge Mass., London

BOLTON, G. E. AND A. OCKENFELS (2000) ERC: a theory of equity, reciprocity and competition, *Amer. Econ. Rev.* 90: 166-193

BOWLES, S. (1998) Endogenous preferences: the cultural consequences of markets and other economic institutions, *J. Econ. Lit.* 36(1): 75-111

CYERT, R. M., DEGROOT, M. H. (1975) Adaptive Utility. In R.H. Day and T. Groves, eds., *Adaptive Economic Models*, p. 223-46, Academic Press, New York, 1975

DEKEL, E., J. ELY AND O. YILANKAYA (2007) Evolution of preferences, *Rev. Econ. Stud.* 74: 685-704

DIETRICH, F. (2008) Anti-terrorism politics and the risk of provoking, *unpublished*

DUFWENBERG, M. AND G. KIRCHSTEIGER (2004) A theory of sequential reciprocity, *Games Econ. Behav.* 47(2): 268-298

FALK, A. AND U. FISCHBACHER (2006) A theory of reciprocity, *Games Econ. Behav.* 54 (2): 293-315

FEHR, E. AND S. GÄCHTER (1998) Reciprocity and economics: the economic implications of homo reciprocans, *Europ. Econ. Rev.* 42: 845-859

GEANAKOPLOS, J., D. PEARCE AND E. STACCHETTI (1989) Psychological Games and Sequential Rationality, *Games Econ. Behav.* 1(1): 60-80

HAMMOND, P. (1976) Changing tastes and coherent dynamic choice, *Rev. Econ. Stud.* 43: 159–173

HÖLDER, O. (1901) Die Axiome der Quantitaet an die Lehre vom Mass, *Ber. ueber d. Verb. d. K. Saechsischen Ges. d. Wiss. zu Leipzig, Math.-Phys. Cl.* 53: 1-64

HUNTINGTON, E. V. (1902a) A complete set of postulates for the theory of absolute continuous magnitude, *Trans. Amer. Math. Soc.* 3: 264-279

HUNTINGTON, E. V. (1902b) A complete set of postulates for the theories of positive integral and positive rational numbers, *Trans. Amer. Math. Soc.* 3: 280-284

Maccheroni, F., Marinacci, M., Rustichini, A. (forthcoming) Social Decision Theory: Choosing within and between Groups, *Review of Economic Studies*

NAKADA, O. (1951) Partially ordered Abelian semigroups I: on the extension of the strong partial order defined on abelian semigroups, *J. Fac. Sci. Hokkaido Univ.* 11: 181-189

O'DONOGHUE, E. AND M. RABIN (1999) Doing it now or doing it later, *Amer. Econ. Rev.* 89: 103-124

POLAK, R. (1976) Interdependent preferences, *Amer. Econ. Rev.* 66(3): 309-20

RABIN, M. (1993) Incorporating fairness into game theory and economics, *Amer. Econ. Rev.* 83: 1281-1302

RABIN, M. (1998) Psychology and economics, *J. Econ. Lit.* 36(1): 11-46

SETHI, R. AND E. SOMANATHAN (2001) Preference evolution and reciprocity, *J. Econ. Theory* 97: 273-297

SETHI, R. AND E. SOMANATHAN (2003) Understanding reciprocity, *J. Econ. Behav. Organ.* 50(1): 1-27

SEN, A. (1977) Rational fools: a critique of the behavioural foundations of economic theory, *Philosophy & Public Affairs* 6(4): 317-344

SEN, A. (1979) Utilitarianism and welfarism, *Journal of Philosophy* 76(9): 463-489

SEN, A. (1985) *Commodities and capabilities*, Amsterdam: North-Holland

STIGLER, G. J. AND G. S. BECKER (1977) De gustibus non est disputandum, *Amer. Econ. Rev.* 67: 76-90

STROTZ, R. H. (1955-56) Myopia and inconsistency in dynamic utility maximization, *Rev. Econ. Stud.* 23(3): 165-180

TAMARI, D. (1949) Groupoïdes reliés et demi-groupes ordonnés, *C. R. Acad. Sci. Paris* 228: 1184-1186

A Appendix: Proofs

A.1 Proof of Theorem 1

The proof of Theorem 1 rests on some preparatory lemmas.

Lemma 1 *Suppose R_1. Composition $\circ$ is an associative operation on $\mathbf{E}$ (hence $(\mathbf{E}, \circ)$ is a semigroup).*

Proof. Assume R_1. Let $e, e', e'' \in \mathbf{E}$. By (1) I have to show for all $c \in \mathbf{C}$ that $c|e(e'e'') = c|(ee')e''$, which holds because, applying the definition of composition and R_1 repeatedly,

$$c|e(e'e'') = c|e|e'e'' = c|e|e'|e'' \text{ and } c|(ee')e'' = c|ee'|e'' = c|e|e'|e''. \blacksquare$$

Lemma 2 *Assume A and R_1. For all experiences e, e', their composition's attractor $\overline{ee'}$ is weakly between $\overline{e}$ and $\overline{e'}$, and $\overline{ee'} \neq \overline{e}$ if $\overline{e} \neq \overline{e'}$.*

As an example for Lemma 2, if states are levels of risk-aversion, experience e attracts to high risk-aversion, and experience e' to low one, then the composition ee' attracts to some intermediate risk-aversion.

Proof. Assume A and R_1, and let $e, e' \in \mathbf{E}$. Consider three cases.

1. First suppose $\bar{e} = \bar{e}'$. Take a $c \in \mathbf{C}$ with this state. Applying twice A and then R_1, $\bar{c} = \overline{c|e} = \overline{c|e|e'} = \overline{c|ee'}$. So $\bar{c} = \overline{c|ee'}$. Hence $\overline{ee'} = \bar{c}$ by A, i.e. $\overline{ee'} = \bar{e} = \bar{e}'$, as desired.

2. Now suppose $\bar{e} < \bar{e}'$. Take $c \in \mathbf{C}$ with state $\bar{c} = \overline{ee'}$. I have to show that $\bar{c} \in (\bar{e}, \bar{e}']$, and do so in two claims.

Claim 1. $\bar{c} \le \bar{e}'$. For a contradiction, let $\bar{c} > \bar{e}'$. Then $\overline{c|e} < \bar{c}$ (by A), and hence $\overline{c|e|e'} < \bar{c}$ (by A and as $\bar{e}', \overline{c|e} < \bar{c}$). But $\overline{c|e|e'} = \overline{c|ee'} = \bar{c}$ (by R_1 and then A), a contradiction. Q.e.d.

Claim 2. $\bar{c} > \bar{e}$. Suppose for a contradiction that $\bar{c} \le \bar{e}$. Then either $\bar{c} = \bar{e}$ or $\bar{c} < \bar{e}$. In the first case, $\overline{c|e} = \bar{c}$, and so (by $\bar{c} < \bar{e}'$ and A) $\overline{c|e|e'} > \bar{c}$. In the second case, $\overline{c|e} > \bar{c}$ and so (by $\bar{e}' > \bar{c}$ and A) $\overline{c|e|e'} > \bar{c}$. So in either case $\overline{c|e|e'} > \bar{c}$. But $\overline{c|e|e'} = \overline{c|ee'} = \bar{c}$ (by R_1 and then A), a contradiction. Q.e.d.

3. Finally, if $\bar{e}' < \bar{e}$, the proof that $\overline{ee'} \in [\bar{e}', \bar{e})$ is analogous to the proof under 2. $\blacksquare$

Lemma 3 *Assume A and R_1. For all constitutions c and experiences e, e',*

(a) $\overline{c|e^n} \to \bar{e}$ *as* $n \to \infty$ *if I_1 holds;*
(b) $\overline{c|e^n e'} \to \bar{e}$ *as* $n \to \infty$ *if I holds.*

Proof. Assume A and R_1 and consider $e, e' \in \mathbf{E}$ and $c \in \mathbf{C}$.

(a) Assume I_1. Consider any constitution $c' \in \mathbf{C}$ with state $\bar{c}' = \bar{e}$. Since by I_1 $\overline{c|e^n} - \overline{c'|e^n} \to 0$ as $n \to \infty$, where by A each $\overline{c'|e^n}$ equals $\bar{e}$, we have $\overline{c|e^n} \to \bar{e}$ as $n \to \infty$.

(b) Assume I. Let $e, e' \in \mathbf{E}$ and $c \in \mathbf{C}$. Since by I_2 $\overline{c|e^n e'} - \overline{c|e^n} \to 0$ as $n \to \infty$, where by part (a) $\overline{c|e^n} \to \bar{e}$, we have $\overline{c|e^n e'} \to \bar{e}$ as $n \to \infty$. $\blacksquare$

Lemma 4 *Assume A and R_1. For all experiences e, e', e'',*

(a) $\overline{e'e^n} \to \bar{e}$ *as* $n \to \infty$ *if I_1 holds;*
(b) $\overline{e^n e'} \to \bar{e}$ *as* $n \to \infty$ *if I holds;*
(c) $\overline{e'e^n e''} \to \bar{e}$ *as* $n \to \infty$ *if I holds.*

Proof. Suppose A and R_1 and let $e, e', e'' \in \mathbf{E}$.

(a) Assume I_1. If $\bar{e} = \bar{e}'$ then by Lemma 2 $\overline{e'e^n} = \bar{e} \to \bar{e}$ as $n \to \infty$. Now let $\bar{e}' < \bar{e}$ (the case $\bar{e}' > \bar{e}$ is analogous). By Lemma 2 (and a simple induction on n) $\overline{e'e^{n+1}} \in [\overline{e'e^n}, \bar{e}]$ for all $n \ge 0$ (where $e'e^0$ stands for e'). So the sequence $(\overline{e'e^n})_{n \ge 0}$ is weakly increasing and upper bounded by $\bar{e}$. Hence $\overline{e'e^n} \to s$ for some $s \le \bar{e}$. As $\mathbf{S}$ is topologically closed, s is in $\mathbf{S}$, i.e. is a state. For a contradiction, assume $s < \bar{e}$. Let c be any constitution with state $\bar{c} = s$. We have $\overline{c|e'e^n} = \overline{c|e'|e^n} \to \bar{e}$ by part (a) of Lemma 3. So there is an $n \ge 0$ such that $\overline{c|e'e^n} > \bar{c}$. However, $\overline{c|e'e^n} \le \bar{c}$ by $\overline{e'e^n} \le \bar{c}$ and A, a contradiction.

(b) Assume I. The case that $\bar{e} = \bar{e}'$ can be treated like in part (a). Now let $\bar{e}' < \bar{e}$ (the case $\bar{e}' > \bar{e}$ is analogous). Like in (a), it can be seen that $(\overline{e^n e'})_{n \ge 0}$

is a weakly increasing sequence converging to some state s. For a contradiction, assume $s < \bar{e}$. Letting c be a constitution of state $\bar{c} = s$, we have $\overline{c|e^n e'} \to \bar{e}$ by part (b) of Lemma 3. So there is an $n \geq 0$ such that $\overline{c|e^n e'} > \bar{c}$ $(= s)$. But $\overline{c|e^n e'} \leq \bar{c}$ by $\overline{e^n e'} \leq \bar{c}$ and A, a contradiction.

(c) Assume I. It suffices to show that (i) $\overline{e' e^{2n} e''} \to \bar{e}$ as $n \to \infty$ and (ii) $\overline{e' e^{2n+1} e''} \to \bar{e}$ as $n \to \infty$. I only show (i), as (ii) follows from (i) by replacing e'' by ee''. By Lemma 2, for all n the state $\overline{e' e^{2n} e''} = \overline{(e' e^n)(e^n e'')}$ is weakly between $\overline{e' e^n}$ and $\overline{e^n e''}$. Hence, as by parts (a) and (b) $\overline{e' e^n}$ and $\overline{e^n e''}$ both converge to $\bar{e}$, so does $\overline{e' e^{2n} e''}$. $\blacksquare$

As most work is contained in the above lemmas, Theorem 1 now follows easily.

Proof of Theorem 1. Assume A, I and R_1, and let $e, e' \in \mathbf{E}$. For all $n \in \{1, 2, ...\}$, Lemma 1 gives the equation $\overline{(ee')^{n+1}} = \overline{e(e'e)^n e'}$, whose left resp. right hand side converges to $\overline{ee'}$ resp. $\overline{e'e}$ as $n \to \infty$ by Lemma 4. So, $\overline{ee'} = \overline{e'e}$. $\blacksquare$

A.2 Proof of Theorem 2

Some lemmas must first be established.

Lemma 5 *Assume R_1.*

  (a) *Constitutions c, c' are identical if $\bar{c} = \bar{c'}$ and $\overline{c|e} = \overline{c'|e}$ for all experiences e.*
  (b) *Experiences e, e' are identical if $\overline{c|e} = \overline{c|e'}$ and $\overline{c|e\tilde{e}} = \overline{c|e'\tilde{e}}$ for all constitutions c and experiences $\tilde{e}$.*

Proof. Assume R_1.

(a) Consider constitutions c, c' such that $\bar{c} = \bar{c'}$ and $\overline{c|.} = \overline{c'|.}$. By (2) and as $\bar{c} = \bar{c'}$, we have $c = c'$ if, for all $e_1, ..., e_k \in \mathbf{E}$ $(k \geq 1)$, $\overline{c|e_1 \cdots |e_k} = \overline{c'|e_1 \cdots |e_k}$. By R_1, the latter is equivalent to $\overline{c|e_1 \cdots e_k} = \overline{c'|e_1 \cdots e_k}$, which holds by $\overline{c|.} = \overline{c'|.}$.

(b) Consider experiences e, e' such that $\overline{c|e} = \overline{c|e'}$ and $\overline{c|e\tilde{e}} = \overline{c|e'\tilde{e}}$ for all constitutions c and experiences $\tilde{e}$. Then, using R_1, $\overline{c|e|\tilde{e}} = \overline{c|e'|\tilde{e}}$ for all constitutions c and experiences $\tilde{e}$. So, by part (a) applied to the constitutions $c|e$ and $c|e'$, we have $c|e = c|e'$ for all constitutions c. This implies $e = e'$ by (1). $\blacksquare$

Lemma 6 *Assume A, I and R_1. If experiences e, e' attract to different states, their composition's attractor $\overline{ee'}$ is strictly between the attractors $\bar{e}$ and $\bar{e'}$.*

Proof. Assume A, I and R_1, and let $e, e' \in \mathbf{E}$ have distinct state. By Lemma 2, $\overline{ee'}$ is weakly between $\bar{e}$ and $\bar{e'}$. Also by Lemma 2, $\overline{ee'} \neq \bar{e}$ and $\overline{e'e} \neq \bar{e'}$, the latter implying $\overline{ee'} \neq \bar{e'}$ by Theorem 1. So $\overline{ee'}$ is strictly between $\bar{e}$ and $\bar{e'}$. $\blacksquare$

We now prove that something somewhat stronger than R_2 holds if we combine R_2 with R_1: *all* constitutions c are reachable from the *same* weak constitution c_w, and this through a *single* experience. Formally:

Lemma 7 *Assume $R_{1,2}$. For every non-weak constitution c there is an experience e_c such that $c = c_w|e_c$ for all weak constitutions c_w.*

Proof. Assume $R_{1,2}$. Let c be a non-weak constitution. By R_2, $c = c_w|e_1|\cdots|e_n$ for some weak constitution c_w and experiences $e_1, ..., e_n$. So, by R_1, $c = c_w|e_c$ where $e_c := e_1 \cdots e_n$. Now let c'_w be an arbitrary weak constitution. I have to show that $c = c'_w|e_c$, i.e. that $c_w|e_c = c'_w|e_c$. By Lemma 5, it suffices to show that (i) $\overline{c_w|e_c} = \overline{c'_w|e_c}$ and (ii) $\overline{c_w|e_c|e} = \overline{c'_w|e_c|e}$ for all experiences e. Equality (i) holds as it reduces to $\overline{e_c} = \overline{e_c}$ by the weakness of c_w and c'_w. The equality in (ii) holds as it reduces to $\overline{c_w|e_c e} = \overline{c'_w|e_c e}$ by R_1, hence to $\overline{e_c e} = \overline{e_c e}$ by the weakness of c_w and c'_w. ∎

Lemma 8 *Assume A, I and $R_{1,2}$. Experiences e_1, e_2 are identical if $\overline{ee_1} = \overline{ee_2}$ for all experiences e.*

Proof. Assume A, I and $R_{1,2}$. Let $e_1, e_2 \in \mathbf{E}$ satisfy $\overline{ee_1} = \overline{ee_2}$ for all $e \in \mathbf{E}$. By part (b) of Lemma 5, it suffices to show the following two claims.

Claim 1. $\overline{c|e_1} = \overline{c|e_2}$ *for all constitutions c.* Let $c \in \mathbf{C}$. Write c as $c = c_w|e$ according to Lemma 7. Using first R_1 and then the weakness of c_w, we have $\overline{c|e_1} = \overline{c_w|ee_1} = \overline{ee_1}$, and similarly $\overline{c|e_2} = \overline{c_w|ee_2} = \overline{ee_2}$. So I have to show that $\overline{ee_1} = \overline{ee_2}$, which holds by assumption on e_1, e_2. Q.e.d.

Claim 2. $\overline{c|e_1\tilde{e}} = \overline{c|e_2\tilde{e}}$ *for all constitutions c and experiences $\tilde{e}$.* Let $c \in \mathbf{C}$ and $\tilde{e} \in \mathbf{E}$. Again write c as $c = c_w|e$ according to Lemma 4. Applying first R_1, then the weakness of c_w, and then Theorem 1, we have $\overline{c|e_1\tilde{e}} = \overline{c_w|ee_1\tilde{e}} = \overline{ee_1\tilde{e}} = \overline{\tilde{e}ee_1}$; and similarly, $\overline{c|e_2\tilde{e}} = \overline{c_w|ee_2\tilde{e}} = \overline{ee_2\tilde{e}} = \overline{\tilde{e}ee_2}$. So I have to show that $\overline{\tilde{e}ee_1} = \overline{\tilde{e}ee_2}$, which holds by assumption on e_1, e_2. ∎

We are now ready to prove Theorem 2.

Proof of Theorem 2. Assume A, I, AC and $R_{1,2}$, and let $e, e' \in \mathbf{E}$. By Lemma 8 it suffices to show $\overline{\hat{e}ee'} = \overline{\hat{e}e'e}$ for all $\hat{e} \in \mathbf{E}$. Let $\hat{e} \in \mathbf{E}$. Consider three exhaustive cases.

Case 1: $\overline{\hat{e}} = \overline{e} = \overline{e'}$. Write s for this state. Applying Lemma 2 repeatedly, we have $s = \overline{\hat{e}e} = \overline{\hat{e}ee'}$, and similarly $s = \overline{\hat{e}e'} = \overline{\hat{e}e'e}$. So $\overline{\hat{e}ee'} = \overline{\hat{e}e'e}$.

Case 2: $\overline{e} \neq \overline{e'}$. Then $\overline{e'} \neq \overline{ee'}$ by Lemma 6. So $\overline{ee'} \neq \overline{e'ee'}$, again by Lemma 6. Moreover, $\overline{e'ee'} = \overline{e'e'e}$ by Theorem 1. So, letting c_w be any weak constitution (it exists by R_2), we have $\overline{c_w|e'ee'} = \overline{c_w|e'e'e}$ by the weakness of c_w, and hence by R_1 $\overline{c|ee'} = \overline{c|e'e}$ where $c := c_w|e'$. By Theorem 1, $\overline{ee'} = \overline{e'e}$. In summary, I have shown that ee' and $e'e$ have a same state – call it s – and that $\overline{c|ee'} = \overline{c|e'e} \neq s$. So, by AC, $\overline{c'|ee'} = \overline{c'|e'e}$ for all $c' \in \mathbf{C}$. Applying this to $c' = c_w|\hat{e}$, I obtain $\overline{c_w|\hat{e}|ee'} = \overline{c_w|\hat{e}|e'e}$, hence by R_1 $\overline{c_w|\hat{e}ee'} = \overline{c_w|\hat{e}e'e}$, and so by the weakness of c_w $\overline{\hat{e}ee'} = \overline{\hat{e}e'e}$.

Case 3: $\overline{e'} \neq \overline{\hat{e}}$. This case reduces to case 2 as by Theorem 1 $\overline{\hat{e}ee'} = \overline{ee'\hat{e}}$ and $\overline{\hat{e}e'e} = \overline{e\hat{e}e'}$. ∎

A.3 Proof of Theorem 3

To prove Theorem 3, I analyse the structure $(\mathbf{E}_s, \geq, \circ)$ (for $s \in \mathbf{S}$) using Hölder's (1901) seminal theorem, which states as follows. Recall that a *(totally) ordered semigroup* is a set X endowed with a linear order $\geq$ and an associative binary operation $\circ$ under which $\geq$ is stable (i.e. such that, for all $x, y, z \in X$, if $x \geq y$ then $x \circ z \geq y \circ z$ and $z \circ x \geq z \circ y$). An ordered semigroup $(X, \geq, \circ)$ is a *(totally) ordered group* if $(X, \circ)$ is a group, *commutative* if $\circ$ is commutative, *cancellative* if $\circ$ is cancellative (i.e. from $xz = yz$ or $zx = zy$ follows $x = y$, for all $x, y, z \in X$), and *semi-divisible* if, for all $x, y \in X$ with $x > y$, y divides x (i.e. $x = ya = by$ for some $a, b \in X$). An element x of the ordered semigroup is an *identity* if $xy = y$ for all y (X contains at most one identity), *weakly positive (weakly negative)* if $xy, yx \geq (\leq)y$ for all $y \in X$, and *strictly positive (strictly negative)* if it is weakly positive (weakly negative) and not an identity.[37] The ordered semigroup is *positively ordered* if each $x \in X$ is weakly positive, and *Archimedean* if for all strictly positive (strictly negative) elements x, y there is an integer $n \geq 1$ such that $x^n \geq y$ ($x^n \leq y$).

Lemma 9 *(Hölder 1901; in part Huntington 1902) Every Archimedean cancellative semi-divisible positively ordered semigroup without identity can be embedded into $((0, \infty), \geq, +)$.*

Before I can apply this result, a number of lemmas must be shown.

Lemma 10 *Assume A and AC.[38] For every state s,*

(a) *the strength relation $\geq$ on $\mathbf{E}_s$ is a weak order;*
(b) *the strength relation $\geq$ on $\mathbf{C}_s$ is a weak order.*

Proof. Assume A and AC. Let $s \in \mathbf{S}$.

(a) On $\mathbf{E}_s$, $\geq$ is obviously transitive ($[e \geq e' \& e' \geq e''] \Rightarrow e \geq e'' \; \forall e, e', e'' \in \mathbf{E}_s$). To show completeness, consider $e, e' \in \mathbf{E}_s$ and suppose $e \not\geq e'$. Then there is an $c \in \mathbf{C}$ such that $\overline{c|e}$ is *not* weakly between s and $\overline{c|e'}$. We have $\overline{c} \neq s$: otherwise $\overline{c|e} = \overline{c|e'} = s$ (as e and e' attract to s). W.l.o.g. suppose $\overline{c} > s$ (the proof is analogous if $\overline{c} < s$). Then, by A and as e and e' attract to s, $\overline{c|e} > s$ and $\overline{c|e'} > s$. Hence, as $\overline{c|e} \notin [s, \overline{c|e'}]$, we have $\overline{c|e} > \overline{c|e'} > s$. So $\overline{c|e'}$ is strictly between s and $\overline{c|e}$. Hence, by AC_1, for *every* constitution c' not of state s, $\overline{c'|e'}$ is strictly between s and $\overline{c'|e}$. So for every constitution c' (whether or not of state s) $\overline{c'|e'}$ is weakly between s and $\overline{c'|e}$. That is, $e' \geq e$, as desired.

---

[37] If X has an identity e, x is weakly positive (weakly negative) if and only if $x \geq (\leq)e$.

[38] In fact, part (a) holds given just A and AC_1, and part (b) holds given just A and AC_2.

(b) On $\mathbf{C}_s$, $\geq$ is again obviously transitive. The proof that $\geq$ is complete is analogous to the completeness proof in (a), with the roles of constitutions and experiences inverted and using AC_2 instead of AC_1. ∎

Lemma 11 *Assume A, I, AC and $R_{1,2}$.*[39] *For every state s,*

(a) *the strength relation $\geq$ on $\mathbf{E}_s$ is a linear order;*
(b) *the strength relation $\geq$ on $\mathbf{C}_s$ is a linear order.*

Proof. Assume A, I, AC and $R_{1,2}$ and let $s \in \mathbf{S}$. By Lemma 10, only anti-symmetry remains to be shown in each part. This is done as follows.

(a) Consider $e, e' \in \mathbf{E}_s$ with $e \equiv e'$; we show that $e = e'$. By Lemma 8 it suffices to show that $\overline{\hat{e}e} = \overline{\hat{e}e'}$ for all $\hat{e} \in \mathbf{E}$. So consider any $\hat{e} \in \mathbf{E}$. Letting c_w be a weak constitution (it exists by R_2), and putting $c := c_w|\hat{e}$, it follows from $e \geq e'$ that $\overline{c|e}$ is weakly between s and $\overline{c|e'}$, and from $e' \geq e$ that $\overline{c|e'}$ is weakly between s and $\overline{c|e}$. So $\overline{c|e'} = \overline{c|e'}$, i.e. $\overline{c_w|\hat{e}|e'} = \overline{c_w|\hat{e}|e'}$. By R_1 $\overline{c_w|\hat{e}e'} = \overline{c_w|\hat{e}e'}$, and so by c_w's weakness $\overline{\hat{e}e'} = \overline{\hat{e}e'}$, as desired.

(b) Consider $c, c' \in \mathbf{C}_s$ such that $c \equiv c'$. Then, for all $e \in \mathbf{E}$ attracting to s', we have from $c \geq c'$ that $\overline{c'|e}$ is weakly between s' and $\overline{c|e}$, *and* from $c' \geq c$ that $\overline{c|e}$ is weakly between s' and $\overline{c'|e}$. So $\overline{c|e} = \overline{c'|e}$ for all $e \in \mathbf{E}$, i.e. $c = c'$ by Lemma 5. ∎

Note the large remaining mathematical gap between the linearity of $(\mathbf{E}_s, \geq)$ (shown in Lemma 11) and the embeddability of $(\mathbf{E}_s, \geq, \circ)$ into $((0, \infty), \geq, +)$ (claimed in Theorem 3). This gap is large not only because of the role of composition $\circ$ but also because many linearly ordered sets (such as sets of higher cardinality than $\mathbb{R}$, the lexicographically ordered set $\mathbb{R}^2$ and many well-ordered sets) are not embeddable into the reals. More work is needed to close this gap.

Lemma 12 *Assume A, I, AC and $R_{1,2}$. For all states s and all experiences $e, e' \in \mathbf{E}_s$, $e > e'$ if and only if, for some experience $\hat{e}$, $\overline{e\hat{e}}$ is strictly between s and $\overline{e'\hat{e}}$.*

Proof. Assume A, I, AC and $R_{1,2}$, and let $s \in \mathbf{S}$ and $e, e' \in \mathbf{E}_s$. Let c_w be a weak constitution (there is one by R_2). First, assume there is $\hat{e} \in \mathbf{E}$ such that $\overline{e\hat{e}}$ is strictly between s and $\overline{e'\hat{e}}$. Then $\overline{c_w|e\hat{e}}$ is strictly between s and $\overline{c_w|e'\hat{e}}$. Hence, $\overline{c|e}$ is strictly between s and $\overline{c|e'}$, where $c := c_w|\hat{e}$. So $e > e'$ by Lemma 11. Second, assume $e > e'$. Then there is a $c \in \mathbf{C}$ such that (*) $\overline{c|e'} > \overline{c|e} \geq s$ or $\overline{c|e'} < \overline{c|e} \leq s$. c is obviously non-weak, so that by Lemma 7 we have $c = c_w|\hat{e}$ for some experience $\hat{e}$. As $\overline{c|e} = \overline{c_w|\hat{e}e} = \overline{e\hat{e}}$ and $\overline{c|e'} = \overline{c_w|\hat{e}e'} = \overline{e'\hat{e}}$, (*) implies that $\overline{e'\hat{e}} > \overline{e\hat{e}} \geq s$ or $\overline{e'\hat{e}} < \overline{e\hat{e}} \leq s$. In these inequalities, I can replace $\geq$ by $>$ and $\leq$ by $<$, by Lemma 6. ∎

---

[39] Part (b) holds given just A, AC_2 and R_1.

Lemma 13 *Assume A, I, AC and $R_{1,2}$.*[40] *The assignment $e \mapsto c_w|e$, where c_w is a fixed weak constitution, does not depend on the choice of c_w and defines*

- *a bijection from $\mathbf{E}$ to $\{c \in \mathbf{C} : c$ is not weak$\}$ and*
- *for each state s an (order-)isomorphism between $(\mathbf{E}_s, \geq)$ and $(\{c \in \mathbf{C}_s : c$ is not weak$\}, \geq)$.*

Proof. Assume A, I, AC and $R_{1,2}$. Let c_w be any weak constitution. First, the assignment does not depend on the choice of c_w, because if c'_w is another weak constitution and $e \in \mathbf{E}$ then $c_w|e = c'_w|e$ by an argument in the proof of Lemma 7. Regarding the first bullet point, surjectivity follows from Lemma 7. To show injectivity, consider distinct $e_1, e_2 \in \mathbf{E}$. By Lemma 8 there is an experience e such that $\overline{ee_1} \neq \overline{ee_2}$, hence by Theorem 1 $\overline{e_1e} \neq \overline{e_2e}$. So, by the weakness of c_w, $\overline{c_w|e_1e} \neq \overline{c_w|e_2e}$. Hence, $c_w|e_1e \neq c_w|e_2e$, and so $c_w|e_1|e \neq c_w|e_2|e$, which implies $c_w|e_1 \neq c_w|e_2$.

Regarding the second bullet point, let us restrict the bijection to $\mathbf{E}_s$ (for some s). The restriction is obviously a bijection onto $\{c \in \mathbf{C}_s : c$ is not weak$\}$. To see that it even is an order-isomorphism, consider any $e_+, e_- \in \mathbf{E}_s$. By Lemma 11, it suffices to show that $e_+ > e_- \Rightarrow c_w|e_+ > c_w|e_-$. Assume $e_+ > e_-$. Then by Lemma 12 there is an experience e such that $\overline{e_+e}$ is strictly between s and $\overline{e_-e}$. So, $\overline{c_w|e_+e}$ is strictly between s and $\overline{c_w|e_-e}$, implying $c_w|e_+ > c_w|e_-$ by Lemma 11. ∎

Lemma 14 *Assume A, I, AC, $R_{1,2}$. For each state s, $(\mathbf{E}_s, \geq, \circ)$ is an Archimedean positively ordered semigroup.*

Proof. Let $s \in \mathbf{S}$. By Lemma 2, $\circ$ indeed defines an operation on $\mathbf{E}_s$.

Claim 1. $(\mathbf{E}_s, \geq, \circ)$ *is an ordered semigroup.* Given Lemmas 1 and 11, $\circ$ is associative and $\geq$ linear. It remains to show stability of $\geq$ under $\circ$. Let $e, \dot{e}, e' \in \mathbf{E}_s$ with $e \geq \dot{e}$. I show $ee' \geq \dot{e}e'$ (which by Theorem 2 also implies $e'e \geq e'\dot{e}$). Assume for a contradiction that $\dot{e}e' > ee'$. By Lemma 12 there is an $\hat{e} \in \mathbf{E}$ such that $\overline{\dot{e}e'\hat{e}}$ is strictly between s and $\overline{ee'\hat{e}}$. This implies, again by Lemma 12, that $\dot{e} > e$. Contradiction. Q.e.d.

Claim 2. $(\mathbf{E}_s, \geq, \circ)$ *is positively ordered.* Assume for a contradiction that $e \in \mathbf{E}_s$ is strictly negative, i.e. $e' > ee' = e'e$ for an $e' \in \mathbf{E}_s$. Then by Lemma 12 there is an $\hat{e} \in \mathbf{E}$ such that $\overline{e'\hat{e}}$ is strictly between $\overline{ee'\hat{e}}$ and $s = \overline{e}$, a contradiction by Lemma 2. Q.e.d.

Claim 3. $(\mathbf{E}_s, \geq, \circ)$ *is Archimedean.* Let $e, \dot{e} \in \mathbf{E}_s$ be strictly positive. I have to find an integer $n \geq 1$ such that $e^n \geq \dot{e}$. If $e \geq \dot{e}$, take $n = 1$. Now suppose $\dot{e} > e$. Then by Lemma 12 there is an $\hat{e} \in \mathbf{E}$ such that $\overline{\dot{e}\hat{e}}$ is strictly between s and $\overline{e\hat{e}}$. By Lemma 4, $\overline{e^n\hat{e}} \to s$ as $n \to \infty$, and so (using that $\overline{\dot{e}\hat{e}} \neq s$ by Lemma

---

[40]Condition AC is not needed for the first bullet point.

6) there is an n such that $\overline{e^n \hat{e}}$ is strictly between s and $\overline{\ddot{e}\hat{e}}$. So, by Lemma 12, $e^n > \dot{e}$. ∎

Lemma 15 *Assume A, I, AC and $R_{1,2}$. For all experiences $e, e', \dot{e}$, if $e\dot{e} = e'\dot{e}$ then $e = e'$ (i.e. ∘ is cancellative).*

Proof. Consider experiences $e, e', \dot{e}$ such that $e\dot{e} = e'\dot{e}$.

Case 1: $\overline{\dot{e}} \neq \overline{e}$. For all $n \geq 1$ we have $e^n \dot{e} = e'^n \dot{e}$ because

$$e^n \dot{e} = e^{n-1}e'\dot{e} = e'e^{n-1}\dot{e} = e'e^{n-2}e'\dot{e} = e'^2 e^{n-2}\dot{e} = ... = e'^n \dot{e}.$$

The state $\overline{e^n \dot{e}} = \overline{e'^n \dot{e}}$ converges to $\overline{e}$ and also to $\overline{e'}$ by Lemma 4. This already gives us $\overline{e} = \overline{e'}$. Now, let c_w be a weak constitution, and consider the constitutions $c := c_w | e$ and $c' := c_w | e'$. Note that $\overline{c} = \overline{c'}$, and that $\dot{e}$ is equally attracted by c as by c', i.e. $\overline{c|\dot{e}} = \overline{c'|\dot{e}} (= \overline{e\dot{e}})$, where this state differs from $\overline{e}$ by Lemma 6. So $c \equiv c'$ by AC, and hence $c = c'$ by Lemma 11. So, by Lemma 13. $e = e'$. Q.e.d.

Case 2: $\overline{\dot{e}} = \overline{e}$. First assume all experiences attract to $\overline{e}$. Then, by Lemma 8, there exists a single experience; hence, $e = e'$, as desired. Now assume there is an experience $\ddot{e}$ attracting to $\overline{\overline{e}} \neq \overline{e}$. Consider the experiences $\tilde{e} := \ddot{e}e$ and $\tilde{e}' := \ddot{e}e'$. We have $\tilde{e}\dot{e} = \tilde{e}'\dot{e}$ (by $\tilde{e}\dot{e} = \ddot{e}e\dot{e} = \ddot{e}e'\dot{e} = \tilde{e}'\dot{e}$), where $\overline{e} \neq \overline{\ddot{e}}$ (by Lemma 6), i.e. $\overline{\dot{e}} \neq \overline{\tilde{e}}$. So, by Case 1 above, $\tilde{e} = \tilde{e}'$, i.e. $\ddot{e}e = \ddot{e}e'$. Noting that $e\ddot{e} = e'\ddot{e}$ (by Theorem 2) with $\overline{\ddot{e}} \neq \overline{e}$, I can again apply Case 1 to infer $e = e'$. ∎

Lemma 16 *Assume A, I and $R_{1,2}$, and let the model be not trivial. Then $(\mathbf{E}, \circ)$ contains no idempotent, i.e. no e with $e^2 = e$. In particular, each $(\mathbf{E}_s, \circ)$ ($s \in \mathbf{S}$) contains no idempotent, hence no identity.*

Proof. Assume A, I, $R_{1,2}$ and non-triviality. Let $e \in \mathbf{E}$.

Claim 1. There is an $e' \in \mathbf{E}$ such that $\overline{e'} \neq \overline{e}$. Suppose the contrary. By Lemma 8, $\mathbf{E} = \{e\}$. Hence, by non-triviality, there is a $c \in \mathbf{C}$ such that (*) $\overline{c|e} \neq \overline{e}$. So c is non-weak, hence by Lemma 13 of the form $c = c_w | e_c$ for some weak $c_w \in \mathbf{C}$ and some $e_c \in \mathbf{E}$. As $\mathbf{E} = \{e\}$, $e_c = e$, whence $\overline{c} = \overline{c_w | e} = \overline{e}$. So $\overline{c|e} = \overline{e}$, contradicting (*). Q.e.d.

Let e' be as in Claim 1. Applying Lemma 6 twice, we have $\overline{e'e} \neq \overline{e}$, and hence $\overline{e'e^2} \neq \overline{e'e}$. So $e^2 \neq e$. ∎

While all but one of Hölder's hypotheses have been shown to hold for our ordered semigroup $(\mathbf{E}_s, \geq, \circ)$ ($s \in \mathbf{S}$), Hölder's semi-divisibility hypothesis need not hold.[41] So Hölder's Theorem cannot be applied directly. To overcome this obstacle, the proof of Theorem 3 will first embed $(\mathbf{E}_s, \geq, \circ)$ into a larger ordered

---

[41] Consider the submodel of the linear model obtained by redefining $\mathbf{E}$ as $\mathbf{S} \times (1, \infty)$ and $\mathbf{C}$ as $\mathbf{S} \times (\{0\} \cup (1, \infty))$. Then $(\mathbf{E}_s, \geq, \circ) \equiv ((1, \infty), \geq, +)$, which is not semi-divisible because $3 > 2$ but there is no $z \in (1, \infty)$ with $3 = 2 + x$.

semigroup, to which Hölder's Theorem can be applied. More precisely, $(\mathbf{E}_s, \geq, \circ)$ is embedded into the positive part of its *ordered group extension*, drawing on another fundamental algebraic result:

Lemma 17 *(Tamari 1949, Alimov 1950, Nakada 1951) For every commutative cancellative ordered semigroup* $(X, \geq, \circ)$,

- *there exists an, up to isomorphism unique, smallest commutative ordered group into which* $(X, \geq, \circ)$ *can be embedded; it is denoted* $(\hat{X}, \geq, \circ)$ *and called the* ordered group extension *of* $(X, \geq, \circ)$;
- $X \subseteq \hat{X}^+$ $(:= \{x \in \hat{X} : x$ *is strictly positive*$\})$ *if* $(X, \geq, \circ)$ *is positively ordered without identity;*
- $(\hat{X}, \geq, \circ)$ *(hence* $(\hat{X}^+, \geq, \circ)$*) is Archimedean if* $(X, \geq, \circ)$ *is Archimedean and positively ordered and contains no* anomalous pair, *i.e. no* x, y *with* $x > y$ *and* $x^n < y^{n+1}$ *for all integers* $n \geq 1$.

For instance, the ordered group extension of $X = \{1, 2, ...\}$ (with $\geq, +$ standardly defined) is $\hat{X} = \{0, \pm 1, \pm 2, ...\}$ (with $\geq, +$ standardly defined). To apply the Tamari-Alimov-Nakada Theorem, a single property must still be shown:

Lemma 18 *Assume A, I, AC and* $R_{1,2}$. *For each state* s, $(\mathbf{E}_s, \geq, \circ)$ *contains no anomalous pair.*

Proof. Assume A, I, AC and $R_{1,2}$. Let $s \in \mathbf{S}$ and $e, \dot{e} \in \mathbf{E}_s$ such that $e > \dot{e}$. By Lemma 12 there is an $\hat{e} \in \mathbf{E}$ such that $\overline{e\hat{e}}$ is strictly between s and $\overline{e'\hat{e}}$. So, since $\overline{(e\hat{e})^n \hat{e}} \to \overline{e\hat{e}}$ as $n \to \infty$ (by Lemma 4), there is an n such that $\overline{(e\hat{e})^n \hat{e}}$ is strictly between s and $\overline{e'\hat{e}}$. In other words, $\overline{e^n \hat{e}^{n+1}}$ is strictly between s and $\overline{e'^{n+1}\hat{e}^{n+1}}$. So, by Lemma 12, $e^n > e'^{n+1}$. ∎

We can now finally prove Theorem 3.

Proof of Theorem 3. Assume A, I, AC and $R_{1,2}$. Suppose the model is not trivial, and let $s \in \mathbf{S}$. By Lemmas 14, 15, 16, 18 and Theorem 2, I may apply the Tamari-Alimov-Nakada Theorem (Lemma 17) to embed the ordered semigroup $(\mathbf{E}_s, \geq, \circ)$ into $(\hat{\mathbf{E}}_s^+, \geq, \circ)$, an Archimedean ordered semigroup. As $(\hat{\mathbf{E}}_s^+, \geq, \circ)$ is moreover semi-divisible, without identity, cancellative and positively ordered (all this by being the strictly positive part of an ordered *group*), it can itself be embedded into $((0, \infty), \geq, +)$ by Hölder's Theorem (Lemma 9). So $(\mathbf{E}_s, \geq, \circ)$ can be embedded into $((0, \infty), \geq, +)$. Hence $(\mathbf{E}_s, \geq, \circ) \equiv (X_s, \geq, +)$ for some set $X_s \subseteq (0, \infty)$ closed under addition.

To show the second bullet point, write $\mathbf{C}_s^* := \{c \in \mathbf{C}_s : c$ is not week$\}$. By Lemma 13, $(\mathbf{C}_s^*, \geq) \equiv (\mathbf{E}_s, \geq)$. So, by the first bullet point $(\mathbf{C}_s^*, \geq) \equiv (X_s, \geq)$. We are done if $\mathbf{C}_s^* = \mathbf{C}_s$, i.e. if $\mathbf{C}_s$ contains no weak constitution. Now suppose it contains one, c_w; then it contains no other one by Lemma 5, and all $c \in \mathbf{C}_s^*$

satisfy $c > c_w$ by definition of (non-)weakness. So, $(\mathbf{C}_s, \geq) = (\mathbf{C}_s^* \cup \{c_w\}, \geq) \equiv (X_s \cup \{0\}, \geq)$. ∎

A.4 Proof of Theorem 4

The proof of the last theorem draws on the following fundamental result of topological algebra due to Arzél (1948) and Tamari (1949). Recall that an ordered semigroup $(X, \geq, \circ)$ is *topological* if its operation $\circ$ is continuous with respect to the order topology on X induced by $\geq$. The notions of 'density' and 'completeness' are to be understood order-theoretically rather than topologically.[42]

Lemma 19 *(Arzél 1948, Tamari 1949) Every cancellative, dense and complete topological ordered semigroup $(X, \geq, \circ)$ with $\#X > 1$ is isomorphic to $(S, \geq, +)$ for some set $S \in \{\mathbb{R}, [0, \infty), (0, \infty), [1, \infty), (1, \infty)\}$ or to the dual $(S, \leq, +)$ thereof.*

To apply this result to the structure $(\mathbf{E}_s, \geq, \circ)$ $(s \in \mathbf{S})$, I now first prove that all premises are satisfied.

Lemma 20 *Assume A, I, AC and $R_{1,2,5}$. For every state s, $\geq$ on $\mathbf{E}_s$ is complete.*

Proof. Assume A, I, AC and $R_{1,2,5}$. Let $s \in \mathbf{S}$. In the definition of completeness, the part on suprema is equivalent to that on infima; so it suffices to show the latter. The claim is obvious if the model is trivial. Now assume it is nontrivial. Let $A \subseteq \mathbf{E}_s$ be a non-empty set that is bounded below, say by $e_< \in \mathbf{E}_s$. I show that A has an infimum in $(\mathbf{E}_s, \geq)$. As $(\mathbf{E}_s, \geq)$ is by Theorem 3 isomorphic to $(X, \geq)$ for some set $X \subseteq (0, \infty)$, there exists a strictly decreasing sequence $(e_k)_{k=1,2,...}$ in A such that for all $e \in A$ we have $e \geq e_k$ for some (sufficiently high) k. It suffices to show that $\{e_k : k = 1, 2, ...\}$ has an infimum (as this infimum is then also one of A).

Claim 1. There is a $c_* \in \mathbf{C}_s$ such that, for all $e \in \mathbf{E}$, $\overline{e_k e}$ converges monotonically to $\overline{c_* | e}$ as $k \to \infty$. For all $e \in \mathbf{E}$, $\overline{e_k e}$ converges (in $\mathbb{R}$): if $\overline{e} = s$ obviously, if $\overline{e} > s$ because $\overline{e_k e}$ is increasing and bounded above by $\overline{e}$, and if $\overline{e} < s$ because $\overline{e_k e}$ is increasing and bounded below by $\overline{e}$. So the sequence of state paths corresponding to the sequence of constitutions $(c_w | e_k)_{k=1,2,...}$, i.e. the sequence of state paths whose k's component is $\overline{(c_w | e_k | e_1 \cdots | e_n)}_{(e_1,...,e_n) \in \cup_{n=0}^{\infty} \mathbf{E}^n} = \overline{(e_k e_1 \cdots e_n)}_{(e_1,...,e_n) \in \cup_{n=0}^{\infty} \mathbf{E}^n}$, converges pointwise. By R_5, the limiting state path is the state path of some $c_* \in \mathbf{C}$. Taking $n = 0$ yields $\overline{e_k} \to \overline{c_*}$, i.e. $s \to \overline{c_*}$, so that $c_* \in \mathbf{C}_s$. Taking $n = 1$ yields, for all $e \in \mathbf{E}$, $\overline{e_k e} \to \overline{c_* | e}$. Q.e.d.

Claim 2. $c_* = c_w | e_*$ for some $e_* \in \mathbf{E}_s$. As the model is not trivial, there is (by an earlier argument) an experience e not of type s. Suppose $\overline{e} > s$ (the proof

[42] An ordered semigroup $(X, \geq, \circ)$ is *dense* if $\geq$ is dense (i.e. for all $x, y \in X$ with $x > y$ there is a $z \in X$ with $x > z > y$), and *complete* if $\geq$ is complete (i.e. every non-empty set $A \subseteq X$ that is bounded below resp. above has an infimum resp. supremum).

is analogous if $\bar{e} < s$. As $e_k \geq e_<$ for all k, $\overline{e_k e} \leq \overline{e_< e}$ for all k. So (by Claim 1) $\overline{c_*|e} \leq \overline{e_< e}$. Hence, $\overline{c_*|e} < \bar{e}$. So $\overline{c_*|e} \neq \bar{e}$. Hence c_* is not weak. Hence, by Lemma 13, $c_* = c_w|e_*$ for some $e_* \in \mathbf{E}_s$. Q.e.d.

Claim 3. e_* is the infimum of $\{e_k : k = 1, 2, ...\}$ (hence of A, completing the proof). First, e_* is a lower bound: each e_k is at least as strong as e_* because, for each $e \in \mathbf{E}$, $\overline{e e_k}$ is (by Claim 1) weakly between s and $\overline{c_*|e}$, i.e. (by Claim 2) weakly between s and $\overline{e_* e}$. Second, consider another lower bound e_{**}, and suppose for a contradiction that $e_{**} > e_*$. Then, by Lemma 12, there is an $e \in \mathbf{E}$ with $\overline{e_{**} e}$ strictly between s and $\overline{e_* e}$. So there is (by Claims 1-2) a k with $\overline{e_{**} e}$ strictly between s and $\overline{e_k e}$. This violates $e_k \geq e_{**}$. ∎

Lemma 21 *Assume A, I, AC and $R_{1,2,4,5}$. For every experience e and state s with $\mathbf{E}_s \neq \emptyset$, $\inf_{e' \in \mathbf{E}_s} |\overline{e e'} - \bar{e}| = 0$.*

Proof. Assume A, I, AC and $R_{1,2,4,5}$ and let $e \in \mathbf{E}$, $s \in \mathbf{S}$ with $\mathbf{E}_s \neq \emptyset$. If $e \in \mathbf{E}_s$ then obviously $\inf_{e' \in \mathbf{E}_s} |\overline{e e'} - \bar{e}| = 0$. Now suppose $e \notin \mathbf{E}_s$; w.l.o.g. let $\bar{e} < s$ (the proof being analogous if $\bar{e} > s$). So I have to show that $\inf_{e' \in \mathbf{E}_s} (\overline{e e'} - \bar{e}) = 0$, i.e. that $\inf_{e' \in \mathbf{E}_s} \overline{e e'} = \bar{e}$.

Claim 1. There is a $c_* \in \mathbf{C}$ such that $\inf_{e' \in \mathbf{E}_s} \overline{e e'} = \overline{c_*|e}$. Consider a sequence $(e_k)_{k=1,2,...}$ in $\mathbf{E}_s$ such that $\overline{e e_k} \to \inf_{e' \in \mathbf{E}_s} \overline{e e'}$ as $k \to \infty$. Like in proof of Claim 1 of the proof of Lemma 20, one can show existence of a $c_* \in \mathbf{C}_s$ such that $\overline{e_k e}$ converges to $\overline{c_*|e}$ as $k \to \infty$; hence $\inf_{e' \in \mathbf{E}_s} \overline{e e'} = \overline{c_*|e}$. Q.e.d.

Claim 2. c_* is weak (hence by Claim 1 $\inf_{e' \in \mathbf{E}_s} \overline{e e'} = \bar{e}$, as desired.) Suppose the contrary. Then, by Lemma 13, there is an $e_* \in \mathbf{E}_s$ such that $c_* = c_w|e_*$ (where c_w is a weak state). We have $\inf_{e' \in \mathbf{E}_s} \overline{e e'} = \overline{c_*|e} = \overline{e e_*}$. So $e' \geq e_*$ for all $e' \in \mathbf{E}_s$, i.e. e_* weakest among all experiences in $\mathbf{E}_s$. Hence, by Lemma 13, $c_w|e_*$ is weakest among all non-weak constitutions in $\mathbf{C}_s$, a violation of R_4. ∎

Lemma 22 *Assume A, I, AC and $R_{1,2,4,5}$. For every state s, if $e \in \mathbf{E}_s$ then $e \geq \hat{e}^2$ for some $\hat{e} \in \mathbf{E}_s$.*

Proof. Assume A, I, AC and $R_{1,2,4,5}$. Let $s \in \mathbf{S}$ and $e \in \mathbf{E}_s$. By R_5 the model is not trivial; hence there exists an $\tilde{e} \in \mathbf{E}$ with $\tilde{e} \notin \mathbf{E}_s$; w.l.o.g. let $\bar{\tilde{e}} < s$ (the proof is analogous if $\bar{\tilde{e}} > 0$). Let $\epsilon := \overline{\tilde{e} e} - \bar{\tilde{e}}$ (> 0). By $\epsilon > 0$ and Lemma 21, there is an $e' \in \mathbf{E}_s$ such that $\overline{\tilde{e} e'} - \bar{\tilde{e}} \leq \epsilon/2$; again by Lemma 21, there is an $e'' \in \mathbf{E}_s$ such that $\overline{\tilde{e} e' e''} - \overline{\tilde{e} e'} \leq \epsilon/2$. It follows that $(\overline{\tilde{e} e'} - \bar{\tilde{e}}) + (\overline{\tilde{e} e' e''} - \overline{\tilde{e} e'}) \leq \epsilon/2 + \epsilon/2$, i.e. $\overline{\tilde{e} e' e''} - \bar{\tilde{e}} \leq \epsilon$. So, letting $\hat{e}$ be the weakest of e' and e'', $\overline{\tilde{e} \hat{e}^2} - \bar{\tilde{e}} \leq \epsilon$, i.e. $\overline{\tilde{e} \hat{e}^2} - \bar{\tilde{e}} \leq \overline{\tilde{e} e} - \bar{\tilde{e}}$. So $\hat{e}^2 \leq e$. ∎

Lemma 23 *Assume A, I, AC and $R_{1,2,4,5}$. For every state s, $\geq$ on $\mathbf{E}_s$ is dense.*

Proof. Assume A, I, AC and $R_{1,2,4,5}$. Let $s \in \mathbf{S}$. If $\mathbf{E}_s = \emptyset$ the claim holds vacuously. Now suppose $\mathbf{E}_s \neq \emptyset$. Let $X_s \subseteq (0, \infty)$ be as in Theorem 3; hence $(\mathbf{E}_s, \geq, \circ) \equiv (X_s, \geq, +)$.

Claim 1. $\inf X_s = 0$ *(where this infimum is formed in* $(\mathbb{R}, \geq)$*, hence exists but needn't belong to* X_s*).* By Lemma 22 there exists a sequence $(e_k)_{k=1,2,\dots}$ in $\mathbf{E}_s$ such that, for all k, $e_{k+1}^2 \leq e_k$. So, by $(\mathbf{E}_s, \geq, \circ) \equiv (X_s, \geq, +)$, there exists a corresponding sequence $(x_k)_{k=1,2,\dots}$ in X_s such that, for all k, $2x_{k+1} \leq x_k$, i.e. $x_{k+1} \leq x_k/2$. In particular, $x_k \to 0$ as $k \to \infty$. Hence $\inf X_s = 0$. Q.e.d.

Claim 2. $(X_s, \geq)$ *is dense (hence* $(\mathbf{E}_s, \geq)$ *is, completing the proof).* Let $x, y \in X_s$ such that $x < y$. By Claim 1, X_s contains a $z < y - x$. Clearly, some multiple nz of z ($n \in \{1, 2, \dots\}$) is strictly between x and y. ∎

Lemma 24 *Assume A, I, AC and* $R_{1,2,4,5}$*. For every state* s*,* $(\mathbf{E}_s, \geq, \circ)$ *is a topological ordered semigroup.*

Proof. Assume A, I, AC and $R_{1,2,4,5}$, and let $s \in \mathbf{S}$. If the model is trivial, the claim is obvious because $\mathbf{E}_s$ is empty or singleton. Now assume non-triviality. By Theorem 3's isomorphism, it suffices to show the claim for the structure $(X, \geq, +)$, where $X := X_x \subseteq (0, \infty)$ is as in Theorem 3. The case $X = \emptyset$ is trivial. Now suppose $X \neq \emptyset$.

Claim 1. X *is topologically dense in* $(0, \infty)$*.* Analogously to Claim 1 in Lemma 23's proof, $\inf X = 0$. This and X's closedness under addition imply the claim. Q.e.d.

Claim 2. The 'intervals' $\{z \in X : x < z < y\}$*,* $x, y \in X$*, form a basis of the order topology.* By definition, every open set is a union of 'intervals' of type (i) $\{z \in X : x < z < y\}$ or (ii) $\{z \in X : x < z\}$ or (iii) $\{z \in X : z < y\}$. Intervals of type (ii) or (iii) are writable as the union of intervals of type (i): $\{z \in X : x < z\} = \cup_{z \in X}\{z \in X : y < x < z\}$ because X has no smallest element (otherwise $\{c \in \mathbf{C}_s : c$ is not weak$\}$ would by Lemma 13 have a smallest element, violating R_4); and $\{z \in X : z < y\} = \cup_{z \in X}\{z \in X : y < x < z\}$ because X has no largest element (as it is closed under addition).

So the intervals of type (i) alone form a basis. Q.e.d.

Claim 3. For all $x, y \in X$ *the inverse* $+^{-1}(\{z \in X : x < z < y\})$ *is open in* X^2 *(which by Claim 2 proves continuity of* $+ : X^2 \to X$*, as desired).* Let $x, y \in X$. It suffices to show that each $(a, b) \in A := +^{-1}(\{z \in X : x < z < y\})$ has an open environment $A_0 \subseteq A$. Let $(a, b) \in A$. So $x < a + b < y$. Hence $\epsilon := \min\{|(a+b) - x|, |(a+b) - y|\} > 0$. By Claim 1, there exist $a_*, a^*, b_*, b^* \in X$ such that $a - \epsilon/2 \leq a_* < a < a^* \leq a + \epsilon/2$ and $b - \epsilon/2 \leq b_* < b < b^* \leq b + \epsilon/2$. The set $A_0 := \{z \in X : a_* < z < a^*\} \times \{z \in X : b_* < z < b^*\}$ contains (a, b), is open in X^2, and is contained in A because all $(a', b') \in A_0$ satisfy $x < a' + b' < y$ by $|(a+b) - (a'+b')| < \epsilon$. ∎

Lemma 25 *Assume A,* $R_{1,2}$*. The model is non-trivial if and only if* $\#\mathbf{E} \geq 2$*.*

Proof. Assume A, $R_{1,2}$. If the model is trivial, then $\#\mathbf{E} \leq 1$ by (1). Now assume non-triviality. Obviously, $\mathbf{E} \neq \emptyset$. Suppose for a contradiction that $\#\mathbf{E} = 1$, say, $\mathbf{E} = \{e\}$.

Claim 1. e is 'fully attracting', i.e., $\overline{c|e} = \bar{e}$ for all $c \in \mathbf{C}$.

Suppose for a contradiction that $c \in \mathbf{C}$ and (*) $\overline{c|e} \neq \bar{e}$. So c is non-weak, hence (by Lemma 7 and by $\mathbf{E} = \{e\}$) of the form $c = c_w|e$ for a (by R_2 existing) weak $c_w \in \mathbf{C}$. Now $\bar{c} = \overline{c_w|e} = \bar{e}$. So $\overline{c|e} = \bar{e}$ by A, contradicting (*). Q.e.d.

Now by non-triviality there are $c, c' \in \mathbf{C}$ such that $c|e \neq c'|e$. So, by (2) there are $n \geq 0$ and $e_1, ..., e_n \in \mathbf{E}$ such that $\overline{c|e|e_1|\cdots|e_n} \neq \overline{c'|e|e_1|\cdots|e_n}$, i.e., by R_2, $\overline{c|ee_1\cdots e_n} \neq \overline{c'|ee_1\cdots e_n}$. As $\mathbf{E} = \{e\}$, $ee_1\cdots e_n = e$, so that $\overline{c|e} \neq \overline{c'|e}$. But $\overline{c|e} = \overline{c'|e} = \bar{e}$ by Clain 1, a contradiction. ∎

I now apply the Arzél-Tamari Theorem to prove the following result.

Lemma 26 *Assume A, I, AC, $R_{1,2,4,5}$ and non-triviality. For every state s to which some experience attracts, $(\mathbf{E}_s, \geq, \circ) \equiv ((0, \infty), \geq, +)$ and $(\mathbf{C}_s, \geq) \equiv ([0, \infty), \geq)$.*

Proof. Assume A, I, AC, $R_{1,2,4,5}$ and non-triviality, and let $s \in \mathbf{S}$ such that $\mathbf{E}_s \neq \emptyset$; hence $\#\mathbf{E}_s = \infty$ by Theorem 3. By Lemmas 15, 20, 23 and 24, $(\mathbf{E}_s, \geq, \circ)$ satisfies all of Arzél-Tamari's premises (Lemma 19), hence is isomorphic to $(S, \geq, +)$ or $(S, \leq, +)$ for some $S \in \{\mathbb{R}, [0, \infty), (0, \infty), [1, \infty), (1, \infty)\}$. But, as $e \circ e > e$ for all $e \in \mathbf{E}_s$, $(\mathbf{E}_s, \geq, \circ)$ is isomorphic to neither $(\mathbb{R}, \geq, +)$, nor $([0, \infty), \geq, +)$, nor $(S, \leq, +)$; and by Lemma 22 it is isomorphic to neither $([1, \infty), \geq, +)$ nor $((1, \infty), \geq, +)$. So $(\mathbf{E}_s, \geq, \circ) \equiv ((0, \infty), \geq, +)$; which by Theorem 3 also implies that $(\mathbf{C}_s, \geq) \equiv ([0, \infty), \geq)$. ∎

Lemma 26 is very useful: one can now define additional structure (operations or relations) on $\mathbf{E}_s$, as long as it is *definable* in terms of the (via Lemma 26 fully understood) structure $\geq, \circ$. Notably, one can define powers of experiences:

Definition 4 *Assume A, I, AC and $R_{1,2,4,5}$. For each experience e and real number $a > 0$, define e^a ('e raised to the power a') as the supremum*

$$e^a := \sup\{e^{\frac{m}{n}} : m, n \in \{1, 2, ...\} \text{ and } \frac{m}{n} \leq a\},$$

where $e^{\frac{m}{n}}$ denotes the experience $e' \in \mathbf{E}_{\bar{e}}$ given by $e'^n = e^m$, i.e. by $\underbrace{e'\cdots e'}_{n} = \underbrace{e\cdots e}_{m}$ (and where existence and uniqueness of the supremum and of $e^{\frac{m}{n}}$ hold by Lemma 26 if the model is non-trivial, and by Lemma 25 otherwise).

Keeping in mind that I use multiplicative notation within $(\mathbf{E}_s, \geq, \circ)$ but additive notation within $((0, \infty), \geq, +)$, raising to the power a in $\mathbf{E}_s$ is the image

under the isomorphism of multiplying by a in $(0, \infty)$.[43] So, the known rules '$(a + b)x = ax + bx$' (distributivity) and '$b(ax) = (ba)x$' (associativity) in $(0, \infty)$ imply by isomorphism the corresponding rules '$e^{a+b} = e^a e^b$' and '$(e^a)^b = e^{ab}$' in $\mathbf{E}_s$. The next lemma contains these two rules and a third (non-obvious) one.

Lemma 27 *Assume A, I, AC, $R_{1,2,4,5}$. For all experiences $e, \dot{e}$ and all reals $a, b > 0$, we have $e^a e^b = e^{a+b}$, $(e^a)^b = e^{ab}$ and $(e\dot{e})^a = e^a \dot{e}^a$.*

Proof. Assume A, I, AC and $R_{1,2,4,5}$. If $\#\mathbf{E} \leq 1$ the claims hold trivially. Now assume $\#\mathbf{E} > 1$. Then the model is non-trivial by Lemma 25, so that the representation of Lemma 26 applies. As mentioned, it remains only to show the third rule '$(e\dot{e})^a = e^a \dot{e}^a$'. Let $e, \dot{e} \in \mathbf{E}$ and $a > 0$.

1. First suppose a is rational, say $a = \frac{m}{n}$ for $m, n \in \{1, 2, ...\}$. Then $(e\dot{e})^a = e^a \dot{e}^a$ because, repeatedly using the rule '$(e^a)^b = e^{ab}$' and commutativity, we have

$$(e\dot{e})^{\frac{m}{n}} = ((e\dot{e})^m)^{\frac{1}{n}} = (e^m \dot{e}^m)^{\frac{1}{n}} = ((e^{\frac{m}{n}})^n (\dot{e}^{\frac{m}{n}})^n)^{\frac{1}{n}} = ((e^{\frac{m}{n}} \dot{e}^{\frac{m}{n}})^n)^{\frac{1}{n}} = e^{\frac{m}{n}} \dot{e}^{\frac{m}{n}}.$$

2. Now let a be arbitrary. Let $s := \overline{e\dot{e}}$ and let $\mathcal{M}$ be the set of all $(m, n) \in \{1, 2, ...\}^2$ such that $\frac{m}{n} \leq a$. I have to show that $e^a \dot{e}^a = \sup_{(m,n) \in \mathcal{M}} (e\dot{e})^{\frac{m}{n}}$, which follows from the following three claims.

Claim 1. $e^a \dot{e}^a \in \mathbf{E}_s$, i.e. $\overline{e^a \dot{e}^a} = s$. It suffices to show that $|\overline{e^a \dot{e}^a} - s| \leq \epsilon$ for all $\epsilon > 0$. Let $\epsilon > 0$. W.l.o.g. suppose $\overline{e} \leq \overline{\dot{e}}$ (the proof is similar else). Then (*) $\overline{e} \leq \overline{e\dot{e}} \leq \overline{\dot{e}}$. Let r be a rational with $0 < r < a$. By $\overline{(e\dot{e})^r} = s$ and Lemma 21 there is a $\delta > 0$ such that for all $b \in (0, \delta)$ we have $|\overline{(e\dot{e})^r e^b} - s| \leq \epsilon$ and $|\overline{(e\dot{e})^r \dot{e}^b} - s| \leq \epsilon$, and hence $\overline{(e\dot{e})^r e^b} - s \geq -\epsilon$ and $\overline{(e\dot{e})^r \dot{e}^b} - s \leq \epsilon$. So, as by (*) $\overline{(e\dot{e})^r e^b} \leq \overline{(e\dot{e})^r e^b \dot{e}^b} \leq \overline{(e\dot{e})^r \dot{e}^b}$, we have $\overline{(e\dot{e})^r e^b \dot{e}^b} - s \geq -\epsilon$ and $\overline{(e\dot{e})^r e^b \dot{e}^b} - s \leq \epsilon$, i.e. $|\overline{(e\dot{e})^r e^b \dot{e}^b} - s| \leq \epsilon$, still for all $b \in (0, \delta)$. Now take any rational $r' > r$ such that $a - \delta \leq r' \leq a$ and choose $b = a - r'$. Note that $e^a \dot{e}^a = e^{r'-r+r+b} \dot{e}^{r'-r+r+b} = e^{r'-r} \dot{e}^{r'-r} e^r \dot{e}^r e^b \dot{e}^b = (e\dot{e})^{r'-r} (e\dot{e})^r e^b \dot{e}^b$, where the last equality holds by part 1. So

$$|\overline{e^a \dot{e}^a} - s| = |\overline{(e\dot{e})^{r'-r}(e\dot{e})^r e^b \dot{e}^b} - s| \leq |\overline{(e\dot{e})^r e^b \dot{e}^b} - s| \leq \epsilon. \text{ Q.e.d.}$$

Claim 2. $e^a \dot{e}^a \geq (e\dot{e})^{\frac{m}{n}}$ for each $(m, n) \in \mathcal{M}$. Let $(m, n) \in \mathcal{M}$. If $\frac{m}{n} = a$ then $e^a \dot{e}^a = (e\dot{e})^{\frac{m}{n}}$ by part 1. If $\frac{m}{n} < a$ then $e^a \dot{e}^a = e^{\frac{m}{n}} e^{a-\frac{m}{n}} \dot{e}^{\frac{m}{n}} \dot{e}^{a-\frac{m}{n}} = (e\dot{e})^{\frac{m}{n}} e^{a-\frac{m}{n}} \dot{e}^{a-\frac{m}{n}} \geq (e\dot{e})^{\frac{m}{n}}$, where the second equality uses part 1. Q.e.d.

Claim 3. No $\tilde{e} \in \mathbf{E}_s$ with $\tilde{e} < e^a \dot{e}^a$ satisfies $\tilde{e} \geq (e\dot{e})^{\frac{m}{n}}$ for all $(m, n) \in \mathcal{M}$. Consider any $\tilde{e} \in \mathbf{E}_s$ with $\tilde{e} < e^a \dot{e}^a$. Then, as $(\mathbf{E}_s, \geq, \circ) = ((0, \infty), \geq, +)$, for sufficiently small $r > 0$ we have $\tilde{e}(e\dot{e})^r < e^a \dot{e}^a$; hence for sufficiently small rational $r > 0$ we have (by part 1) $\tilde{e} e^r \dot{e}^r < e^a \dot{e}^a = e^{a-r} \dot{e}^{a-r} e^r \dot{e}^r$, which (by cancellation) implies $\tilde{e} < e^{a-r} \dot{e}^{a-r}$. Take any $(m, n) \in \mathcal{M}$ with $\frac{m}{n} > a - r$. Then $\tilde{e} < e^{a-r} \dot{e}^{a-r} < e^{a-r} \dot{e}^{a-r} e^{\frac{m}{n}-(a-r)} \dot{e}^{\frac{m}{n}-(a-r)} = e^{\frac{m}{n}} \dot{e}^{\frac{m}{n}}$. ∎

---

[43] Because $ax = \sup\{\frac{m}{n}x : m, n \in \{1, 2, ...\}$ and $\frac{m}{n} \leq a\}$ where $\frac{m}{n}x = x' \in (0, \infty)$ defined by $nx' = mx$.

Lemma 28 *Assume A, I, AC and $R_{1,2,4,5}$. For all experiences e_0, e_1 with $\overline{e_0} < \overline{e_1}$, the assignment $a \mapsto \overline{e_0 e_1^a}$ defines an increasing bijection from $(0, \infty)$ to $(\overline{e_0}, \overline{e_1})$. In particular, $\{\overline{e} : e \in \mathbf{E}\}$ $(\subseteq \mathbf{S})$ is an interval.*

Proof. Assume A, I, AC and $R_{1,2,4,5}$, let $e_0, e_1 \in \mathbf{E}$ with $\overline{e_0} < \overline{e_1}$, and let $f : (0, \infty) \to \mathbb{R}, a \mapsto \overline{e_0 e_1^a}$. Claims 1 and 3 below establish the result.

Claim 1. f is strictly increasing. For all $0 < a < b$, $f(b) = \overline{e_0 e_1^b} = \overline{e_0 e_1^a e_1^{b-a}} > \overline{e_0 e_1^a} = f(a)$, where the inequality holds by $\overline{e_1^{b-a}} > \overline{e_0 e_1^a}$. Q.e.d.

Claim 2. $\lim_{a \to \infty} f(a) = \overline{e_1}$ and $\lim_{a \to 0} f(a) = \overline{e_0}$. By Lemma 4, $f(n) = \overline{e_0 e_1^n} \to \overline{e_1}$ as the natural number n tends to ∞. So (using Claim 1) $f(a) \to \overline{e_1}$ as $a \to \infty$. By a similar argument, $\overline{e_0^b e_1} \to \overline{e_0}$ as $b \to \infty$. So, as by Lemma 27 $\overline{e_0^b e_1} = \overline{(e_0^b e_1)^{1/b}} = \overline{e_0 e_1^{1/b}} = f(\frac{1}{b})$, we have $f(\frac{1}{b}) \to \overline{e_0}$ as $b \to \infty$, i.e. $f(a) \to \overline{e_0}$ as $a \to 0$. Q.e.d.

Claim 3. $f((0, \infty)) = (\overline{e_0}, \overline{e_1})$. Let $s \in (\overline{e_0}, \overline{e_1})$. I show that $f(a^*) = s$ for some $a^* \in (0, \infty)$. We have $\sup f^{-1}((\overline{e_0}, s]) = \inf f^{-1}([s, \overline{e_1}))$, by Claim 1 and $f^{-1}((\overline{e_0}, s]) \cup f^{-1}([s, \overline{e_1})) = (0, \infty)$. Let $a^* := \sup f^{-1}((\overline{e_0}, s]) = \inf f^{-1}([s, \overline{e_1}))$ ($\in [0, \infty]$). Note that $a^* \notin \{0, \infty\}$, because otherwise $f^{-1}((\overline{e_0}, s]) = \emptyset$ or $f^{-1}([s, \overline{e_1})) = \emptyset$, violating Claim 2. So $a^* \in (0, \infty)$. The proof is completed by showing that $f(a^*) = s$.

I first show $f(a^*) \leq s$. For all $n \in \{1, 2, ...\}$, by Lemma 27 $f(\frac{na^*}{n+1}) = \overline{e_0 e_1^{na^*/(n+1)}} = \overline{e_0^{n+1} e_1^{na^*}} = \overline{e_0 (e_0 e_1^{a^*})^n}$. So $f(\frac{na^*}{n+1}) \to \overline{e_0 e_1^{a^*}} = f(a^*)$ as $n \to \infty$ by Lemma 4. As for all n, $f(\frac{na^*}{n+1}) < s$ (by $\frac{na^*}{n+1} < a^* = \inf f^{-1}([s, \infty))$), in the limit $f(a^*) \leq s$.

I finally show $f(a^*) \geq s$. For all $n \in \{1, 2, ...\}$, by Lemma 27 $f(\frac{(n+1)a^*}{n}) = \overline{e_0 e_1^{(n+1)a^*/n}} = \overline{e_0^n e_1^{(n+1)a^*}} = \overline{e_1^{a^*} (e_0 e_1^{a^*})^n}$. So $f(\frac{(n+1)a^*}{n}) \to \overline{e_0 e_1^{a^*}} = f(a^*)$ as $n \to \infty$ by Lemma 4. For all n. $f(\frac{(n+1)a^*}{n}) > s$ (by $\frac{(n+1)a^*}{n} > a^* = \sup f^{-1}((0, s]))$, whence in the limit $f(a^*) \geq s$. ∎

Lemma 29 *Assume $R_{1,3}$. We have $\{\overline{e} : e \in \mathbf{E}\} = \mathbf{S}$.*

Proof. Assume $R_{1,3}$. Clearly, $\{\overline{e} : e \in \mathbf{E}\} \subseteq \mathbf{S}$. Now let $s \in \mathbf{S}$. Choose any $c \in \mathbf{C}_s$. By R_3 there is an $e \in \mathbf{E}$ such that $\overline{c|e} = s$, and by R_1 it follows that $\overline{e} = s$. ∎

Drawing on these lemmas, we can attack the proof of Theorem 4.

Proof of Theorem 4. First, the linear model for a state set $\mathbf{S}$ obviously satisfies all of properties A, I, AC and R_{1-5}; and so do its isomorphic variants, because reparameterisations preserve these properties (in the case of I because an increasing bijection between two state sets is automatically continuous).

Second, I consider a change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\overline{\cdot}), (.|.))$ satisfying A, I, AC and R_{1-5}, and show that it is a reparametrisation of the linear model, to be denoted

$(\mathbf{S}, \hat{\mathbf{E}}, \hat{\mathbf{C}}, (\hat{.}), (.|\hat{.}))$. Specifically, I first define three transformations $\sigma : \mathbf{S} \to \mathbf{S}$, $\epsilon : \hat{\mathbf{E}} \to \mathbf{E}$, $\chi : \hat{\mathbf{C}} \to \hat{\mathbf{C}}$, and then prove in several claims that they define a reparametrisation in the required sense.

W.l.o.g. let $\mathbf{S} = [0, 1]$. (The proof is analogous for other choices of $\mathbf{S}$.) For all states s, let $w(s)$ the (by Lemma 26 uniquely existing) weak constitution in $\mathbf{C}_s$. Further, let $\tilde{w} \in \mathbf{C}$ be an arbitrary weak constitution. Fix experiences $e_0 \in \mathbf{E}_0$ and $e_1 \in \mathbf{E}_1$ (they exist by Lemma 29).

Let any experience $s_x \in \hat{\mathbf{E}} = [0, 1] \times (0, \infty)$ be transformed into $\epsilon(s_x) = e_1^{xs} e_0^{x(1-s)} \in \mathbf{E}$, where the last expression is to be read as e_1^x if $s = 1$ and as e_0^x if $s = 0$. Let any state $s \in \hat{\mathbf{S}} = [0, 1]$ be transformed into $\sigma(s) = \overline{\epsilon(s_1)}$ $(= \overline{e_1^s e_0^{1-s}} \in \mathbf{S} = [0, 1])$. And let each constitution $s_x \in \hat{\mathbf{C}}$ be transformed into

$$
\chi(s_x) = \begin{cases} \tilde{w}|\epsilon(s_x) & \text{if } x > 0 \text{ (i.e. if } s_x \in \hat{\mathbf{E}}) \\ w(\sigma(s)) & \text{if } x = 0 \text{ (i.e. if } s_x \notin \hat{\mathbf{E}}). \end{cases}
$$

Claim 1. $\sigma : [0, 1] \to [0, 1]$ *is a strictly increasing bijection (hence is continuous).* For all $s \in (0, 1)$ we have $\sigma(s) = \overline{e_1^s e_0^{1-s}} = \overline{e_1^{s/(1-s)} e_0}$. So σ is on $(0, 1)$ the composition of the strictly increasing bijection $(0, 1) \to (0, \infty)$, $s \mapsto \frac{s}{1-s}$ and the function $(0, \infty) \to (0, 1), a \mapsto \overline{e_1^a e_0}$, which is also a strictly increasing bijection by Lemma 28. So σ defines a strictly increasing bijection from $(0, 1)$ to $(0, 1)$. σ extends to a strictly increasing bijection from $[0, 1]$ to $[0, 1]$ because $\sigma(0) = \overline{e_0} = 0$ and $\sigma(1) = \overline{e_1} = 1$. Q.e.d.

Claim 2. $\epsilon : \hat{\mathbf{E}} \to \mathbf{E}$ *is bijective.* To show injectivity, consider distinct $s_x, s'_{x'} \in \hat{\mathbf{E}}$.

Case 1: $s = s'$. Then $x \neq x'$. We have $\epsilon(s_x) = e_1^{xs} e_0^{x(1-s)} \neq [e_1^{xs} e_0^{x(1-s)}]^{x'/x} = e_1^{x's} e_0^{x'(1-s)} = \epsilon(s_{x'})$, as desired.

Case 2: $s \neq s'$. Suppose w.l.o.g. that $s < s'$ (the proof is analogous if $s > s'$). We have $\overline{\epsilon(s_x)} = \overline{[\epsilon(s_x)]^{1/x}} = \overline{\left[e_1^{xs} e_0^{x(1-s)} \right]^{1/x}} = \overline{e_1^s e_0^{1-s}}$, and analogously $\overline{\epsilon(s'_{x'})} = \overline{e_1^{s'} e_0^{1-s'}}$. So it suffices to show that $\overline{e_1^s e_0^{1-s}} \neq \overline{e_1^{s'} e_0^{1-s'}}$. This follows from

$$
\overline{e_1^s e_0^{1-s}} < \overline{e_1^{s'-s} e_1^s e_0^{1-s}} = \overline{e_1^{s'} e_0^{1-s}} = \overline{e_1^{s'} e_0^{1-s'} e_0^{s'-s}} < \overline{e_1^s e_0^{1-s'}}.
$$

To show that ϵ is also surjective, consider any $e \in \mathbf{E}$. As σ is bijective (Claim 1), there is and $s \in [0, 1]$ such that $\bar{e} = \sigma(s) = \overline{\epsilon(s_1)}$. As e and $\epsilon(s_1)$ both belong to $\mathbf{E}_{\bar{e}}$, there is (by Lemma 26) an $x > 0$ such that $e = [\epsilon(s_1)]^x$. So

$$
e = [\epsilon(s_1)]^x = \left[e_1^s e_0^{1-s} \right]^x = e_1^{xs} e_0^{x(1-x)} = \epsilon(s_x). \text{ Q.e.d.}
$$

Claim 3. $\chi : \hat{\mathbf{C}} \to \mathbf{C}$ *is bijective.* Note that $\hat{\mathbf{C}}$ is the disjoint union of $\hat{\mathbf{E}}$ and $\{s_0 : s \in [0, 1]\}$ (containing the non-weak resp. weak constitutions). So the claim follows from two observations. First, the restriction $\chi|_{\hat{\mathbf{E}}}$ is bijective between $\hat{\mathbf{E}}$ and $\{c \in \mathbf{C} : c \text{ is not weak}\}$, because it is the composition of the (by Claim 2 bijective)

41

mapping $\epsilon : \hat{\mathbf{E}} \to \mathbf{E}$ and the (by Lemma 13 bijective) mapping $e \mapsto \tilde{w}|e$ from $\mathbf{E}$ to $\{c \in \mathbf{C} : c \text{ is not weak}\}$. Second, the restriction $\chi|_{\{s_0 : s \in [0,1]\}}$ is bijective between $\{s_0 : s \in [0,1]\}$ and $\{c \in \mathbf{C} : c \text{ is weak}\}$ because it is given by the assignment $s_0 \mapsto w(\sigma(s))$, where σ is (by Claim 1) bijective from $[0,1]$ to $[0,1]$. Q.e.d.

Claim 4. $(\bar{\cdot})$ *is the image of* $(\hat{\cdot})$, *i.e.* $\sigma(\hat{c}) = \overline{\chi(c)}$ *for all* $c \in \hat{\mathbf{C}}$. Consider any $s_x \in \hat{\mathbf{C}}$. I have to show that $\overline{\chi(s_x)} = \sigma(\hat{s_x})$, i.e. that $\overline{\chi(s_x)} = \sigma(s)$. If $x = 0$ this holds because $\overline{\chi(s_0)} = \overline{w(\sigma(s))} = \sigma(s)$. If $x > 0$ then it holds because

$$\overline{\chi(s_x)} = \overline{\tilde{w}|\epsilon(s_x)} = \overline{\epsilon(s_x)} = \overline{e_1^{xs} e_0^{x(1-s)}} = \overline{(e_1^s e_0^{1-s})^x} = \overline{e_1^s e_0^{1-s}} = \overline{\epsilon(s_1)} = \sigma(s),$$

where the second equality holds by the weakness of $\tilde{w}$, the fourth by Lemma 27, and all others by definition. Q.e.d.

Claim 5. $(.|.)$ *is the image of* $(.\hat{|}.)$. Consider any $s_x \in \hat{\mathbf{C}}$ and $s'_{x'} \in \hat{\mathbf{E}}$. I have to show that $\chi(s_x)|\epsilon(s'_{x'}) = \chi(s_x \hat{|} s'_{x'})$.

Case 1: $x = 0$ (i.e. s_x is weak). Then $\chi(s_x)|\epsilon(s'_{x'}) = \chi(s_x \hat{|} s'_{x'})$ as, by definition of χ and $(.\hat{|}.)$ and by the weakness of $w(\sigma(s))$,

$$\chi(s_x)|\epsilon(s'_{x'}) = w(\sigma(s))|\epsilon(s'_{x'}) = \tilde{w}|\epsilon(s'_{x'}) \text{ and } \chi(s_x \hat{|} s'_{x'}) = \chi(s'_{x'}) = \tilde{w}|\epsilon(s'_{x'}).$$

Case 2: $x > 0$ (i.e. s_x is not weak). Then, by definition of χ,

$$\chi(s_x)|\epsilon(s'_{x'}) = \tilde{w}|\epsilon(s_x)\epsilon(s'_{x'}) \text{ and } \chi(s_x \hat{|} s'_{x'}) = \tilde{w}|\epsilon(s_x \hat{|} s'_{x'}).$$

So I have to show that $\tilde{w}|\epsilon(s_x)\epsilon(s'_{x'}) = \tilde{w}|\epsilon(s_x \hat{|} s'_{x'})$, i.e. by Lemma 13 that $\epsilon(s_x)\epsilon(s'_{x'}) = \epsilon(s_x \hat{|} s'_{x'})$. The latter holds because, by definition of ϵ and Lemma 27,

$$
\begin{aligned}
\epsilon(s_x)\epsilon(s'_{x'}) &= e_1^{xs} e_0^{x(1-s)} e_1^{x's'} e_0^{x'(1-s')} = e_1^{xs+x's'} e_0^{x(1-s)+x'(1-s')} \\
&= e_1^{xs+x's'} e_0^{x+x'-sx-s'x'}, \\
\epsilon(s_x \hat{|} s'_{x'}) &= \epsilon\left(\left[\frac{xs+x's'}{x+x'}\right]_{x+x'}\right) = e_1^{xs+x's'} e_0^{(x+x')(1-\frac{xs+x's'}{x+x'})} \\
&= e_1^{xs+x's'} e_0^{x+x'-xs-x's'}. \blacksquare
\end{aligned}
$$

A.5   Proof of Theorems 1*-4*

I now sketch the proofs of Theorems 1*-4* as corollaries of Theorems 1-4; full proofs are available on request from the author. Consider a fixed number of dimensions $n \geq 1$ and an n-dimensional change model $\mathcal{M} = (\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ satisfying A. Recall from Section 8 that $\mathcal{M}$ induces a sub-model $\mathcal{M}_{[a,b]}$ for each (sub-) state space of the form of a line segment $[a,b] \subseteq \mathbf{S}$ (where a and b are any distinct states in $\mathbf{S}$), and that Theorems 1-4 continue to apply to such essentially (rather than properly) one-dimensional state spaces.

To prove Theorem 1*, assume $\mathcal{M}$ satisfies this theorem's three axioms and consider experiences $e, e' \in \mathbf{E}$. Choose any essentially one-dimensional (sub-)

42

state space $[a, b] \subseteq \mathbf{S}$ ($a \neq b$) containing $\overline{e}$ and $\overline{e'}$. (One may take $a = \overline{e}$ and $b = \overline{e'}$ as long as $\overline{e} \neq \overline{e'}$.) As can show, the sub-model $\mathcal{M}_{[a,b]}$ inherits the three axioms from $\mathcal{M}$, so that, by Theorem 1, $\overline{ee'} = \overline{e'e}$.[44] This proves Theorem 1*.

The derivation of Theorem 2* is analogous to that of Theorem 1*.[45]

To prove Theorem 3*, assume $\mathcal{M}$ satisfies its conditions, and consider a state $s \in \mathbf{S}$. Choose a $b \in \mathbf{S} \backslash \{s\}$. The sub-model $\mathcal{M}_{[s,b]}$ again inherits all conditions imposed on $\mathcal{M}$, so that Theorem 3 applies to it. Hence, there exists a set $X_s \subseteq (0, \infty)$ closed under addition such that (i) $(\mathbf{E}_s, \geq, \circ)$ is isomorphic to $(X_s, \geq, +)$, and (ii) $(\mathbf{C}_s, \geq)$ is isomorphic to $(X_s, \geq)$ if $\mathbf{C}_s$ contains no weak constitution, and to $(X_s \cup \{0\}, \geq)$ otherwise. Here, the relation $\geq$, the operation $\circ$, and the notion of 'weakness' are initially defined w.r.t. the sub-model $\mathcal{M}_{[s,t],t}$; but one may show that they can equivalently be defined w.r.t. to the original model $\mathcal{M}$. This proves Theorem 3*.[46]

As for Theorem 4*, one easily checks that if $\mathcal{M}$ is isomorphic to the linear model, then all axioms hold. Now suppose the axioms hold. If $n = 1$, then Theorem 4 applies directly, i.e. $\mathcal{M}$ is the linear model up to isomorphism. Now suppose $n > 1$. For any essentially one-dimensional (sub-) state space $[a, b] \subseteq \mathbf{S}$ ($a \neq b$), the sub-model $\mathcal{M}_{[a,b]}$ again inherits all axioms from $\mathcal{M}$, and hence is linear up to isomorphism by Theorem 4. It needs to be verified that the isomorphisms for the various sub-models are 'compatible' in such a way that the full model $\mathcal{M}$ is also linear up to isomorphism. I here only define the transformation through which $\mathcal{M}$ becomes linear, leaving it to the reader to verify that this transformation is indeed an isomorphism to the linear model. Call two experiences $e, f \in \mathbf{E}$ *equally strong* if their composition ef is half way between e and f in terms of the state of attraction: $\overline{ef} = \frac{1}{2}\overline{e} + \frac{1}{2}\overline{f}$. Fix an experience $g \in \mathbf{E}$; it will be used for normalization purposes and is said to have 'strength one'. For each state $s \in \mathbf{S} \backslash \{\overline{g}\}$, there is a unique experience in $\mathbf{E}_{\overline{s}}$ which is equally strong as g, and is thus also said to have 'strength one'; this follows from the linearity up to isomorphism of $\mathcal{M}_{[\overline{g},s]}$. Now define the *strength* of a constitution $c \in \mathbf{C}$ as the

---

[44] To be precise, the sub-model $\mathcal{M}_{[a,b]}$ may violate the conventions (1) and (2), in which case $\mathcal{M}_{[a,b]}$ is not a proper change model. Fortunately though, Theorem 1 immediately extends to such 'improper' models.

[45] I note two technicalities. First, to ensure the sub-model $\mathcal{M}_{[a,b]}$ inherits the conventions (1) and (2), one should choose $[a, b]$ such that $\#([a, b] \cap \{\overline{e} : e \in \mathbf{E}\}) \geq 2$ (which is possible after assuming that $\#\mathbf{E} > 1$ – an unproblematic assumption since the theorem holds trivially if $\#\mathbf{E} \leq 1$). Second, the sub-model may initially fail to inherit one axiom, namely R_2, since the sub-model may fail to contain a weak constitution. To enforce R_2, one should slightly amend the sub-model: artifically add an extra constitution and extend the revision rule such as to render this constitution weak. The same amendment must also be done to the sub-models considered in the proofs of Theorems 3* and 4*.

[46] The given argument requires that we choose b from $\{\overline{e} : e \in \mathbf{E}\}$ (which is always possible) and that $\mathbf{E}_s \neq \varnothing$. The case $\mathbf{E}_s = \varnothing$ is trivial since it suffices to define X_s as $\varnothing$.

degree to which it is unaffected by experiences h of strength one, i.e. as the ratio

$$\sigma(c) := \frac{\left\|\overline{c|h} - \overline{h}\right\|}{\left\|\overline{c|h} - \overline{c}\right\|} \quad \left(= \frac{\text{'distance of posterior state to attractor'}}{\text{'distance of posterior state to prior state'}} \right),$$

where $\|\cdot\|$ denotes the Euclidean norm and where this definition does not depend on the choice of h, as one may show.[47] Similarly, define the *strength* of an experience $e \in \mathbf{E}$ as the degree to e is unaffected by adding an experience h of strength one, i.e. as the ratio

$$\sigma(e) := \frac{\left\|\overline{eh} - \overline{h}\right\|}{\left\|\overline{eh} - \overline{e}\right\|} \quad \left(= \frac{\text{'distance of new attractor to attractor of } h\text{'}}{\text{'distance of new attractor to old attractor'}} \right),$$

where this definition does again not depend on the choice of h.[48] Now, as one can show, the following transformation defines an isomorphism from $\mathcal{M}$ to the linear model with state space $\mathbf{S}$:

- map each state $s \in \mathbf{S}$ to the state of the linear model identical to s,
- map each constitution $c \in \mathbf{C}$ to the constitution s_x of the linear model given by $s = \overline{c}$ and $x = \sigma(c)$,
- map each experience $e \in \mathbf{E}$ to the experience s_x of the linear model given by $s = \overline{e}$ and $x = \sigma(e)$.

---

[47]But it is essential that h has strength one, and that $\overline{h} \neq \overline{c}$ to ensure that $\overline{c|h} \neq \overline{c}$.

[48]It is essential that h has strength one, and that $\overline{h} \neq \overline{e}$ to ensure that $\overline{eh} \neq \overline{e}$.