Franz Dietrich

Modelling change in individual characteristics:
an axiomatic framework

RM/08/045

JEL code: C7, D0, D6, D9

METE⌒⌒R

Modelling change in individual characteristics: an axiomatic framework

Franz Dietrich[1]

November 2008

Abstract

Economic models describe individuals in terms of underlying characteristics, such as taste for some good, sympathy level for another player, time discount rate, risk attitude, and so on. In real life, such characteristics change through experiences: taste for Mozart changes through listening to it, sympathy for another player through observing his moves, and so on. Models typically ignore change, not just for simplicity but also because it is unclear how to incorporate change. I introduce a general axiomatic framework for defining, analysing and comparing rival models of change. I show that seemingly basic postulates on modelling change together 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 solid axiomatic grounds and enabling non-arbitrary incorporation of change into economic models.

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 that, 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 a_t depends on time t or on 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 here 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).

Allowing agent characteristics to change is perfectly compatible with the orthodox methodology of keeping stable preferences (over full histories), and also with assuming stable separable intra-period utilities of intra-period *extended* outcomes, i.e. of pairs (a_t, s) of the outcome a_t in the current period t and the agent's current state s (e.g. his current taste for wine or feelings for someone), where s may depend on time t or on past outcomes. Indeed, Becker is a vigorous defender not only of individual change but also of stable preferences. On the other hand, the present analysis of change in characteristics *could* be used to model preference change (dynamic inconsistency), namely by defining the individual's current preferences as a function of some current characteristic (e.g. a current addiction level or impatience level). I indeed believe that preference change (if of interest) is best explained and represented by change in a well-defined characteristic; hence the relevance of this paper also to modelling dynamic inconsistency. Dynamic inconsistency is popular in empirical and behavioural economics, and unorthodox in theoretical economics. This paper stays neutral on this methodological issue.

The problem so far has not been that change in individual characteristics is seriously denied, but that there is little general theoretical understanding of how to model such change. Even the greatest proponents of individual change and endogenous characteristics, such as Becker and Sen or advocates of dynamic inconsistency, leave many questions open when it comes to concretely modelling change.[2] Allowing individual characteristics to change without at the same time

---

[2]Becker, for instance, distinguishes between whether marginal utility from a good depends positively or negatively on past consumption of this good. This gives interesting qualitative

restricting the way in which they can change runs the risk of a theory with little predictive power: there will likely be many free parameters that are hard to estimate, up to the point of empirical emptiness (i.e. compatibility with every behavioural pattern). The only hope to break the underdetermination and render change and endogeneity more popular in modelling practice is to narrow down the set of possible change patterns, ideally leaving few and easily interpretable parameters. But how should we narrow down the set of possibilities? An axiomatic approach (as pursued here) is one candidate.

Surprisingly, mathematical economics has so far not engaged in a formal axiomatic treatment of individual change, in spite of having driven rational choice axiomatics into high spheres of specialisation and sophistication. It appears long due to take a systematic axiomatic approach to individual change as well. The neglect that change has received is surely not due to a lack of relevance, but perhaps to an implicit assumption that change falls outside the scope of theory and axiomatics. As I hope to show, theory has very much to say on change, and axiomatics is not only possible but even fruitful.

The axioms on change behaviour introduced here have some flavour of 'rationality' postulates, though it is clear that principles of rationality alone cannot tell us how exactly someone's feelings or his taste for wine will (should?) change in the face of certain experiences. The combination of my postulates turns out to severely constrain and 'discipline' change, forcing it to take a simple and convenient form potentially suited for modelling practice: 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

_____

insights but is far from a complete parametric model of change.

3

and Ockenfels (2000), Sethi and Somanathan (2001) and (2003), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006) and Dietrich (2008); on endogenously changing fundamental preferences (i.e. dynamic inconsistency), see Strotz (1955-56), Hammond (1976), and O'Donoghue and Rabin (1999); on preference evolution, see Dekel et al. (2007). Such models of endogeneity have implications for the change behaviour of the endogenous characteristic in question.[3] But they are not (and were not intended as) full-fledged change models, for a variety of reasons.[4] This paper aims at 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. By contrast, this paper's framework is entirely general (but becomes context-specific by suitably specifying its components and parameters).

My theoretical approach is by no means a replacement for empirical research on individual change. Rather, the two approaches are complementary and even need each other, for the following reasons. On the one hand, any model of change that a theoretical approach can recommend will certainly have free parameters whose exact values can only be approximated statistically. For instance, the *linear* change model studied below describes the effect of Peter's unfaithful behaviour on Ann's feelings for him by two parameters: the (presumably low) level to which her feelings are 'attracted' by his bad behaviour, and the 'strength' of this attraction. Calibrating these parameters is an empirical matter. On the other hand, any empirical study of how an individual characteristic changes under some influences presupposes the theoretical step of first choosing a model within which hypotheses can be formulated and tested and free parameters can be estimated. The model should not be chosen ad hoc but should ideally emerge from a transparent axiomatic approach, notably because the phenomenon to be modelled (individual change) happens to be one for which we lack definite prior intuitions.[5]

---

[3] According to a model that makes some individual characteristic X depend on some context (factors) Y, say via a deterministic relationship $X = f(Y)$, each context y (a value of Y) lets the characteristic change from its original value to the value $x = f(y)$. A note of caution: one should be careful with reading more into this 'one-shot' model. In particular, one should not easily infer that a second change of context, say from y to y', will let the characteristic change from $f(y)$ to $f(y')$, because the past experience of context y might have affected the individual's internal constitution, letting him respond differently to new contexts in a 'second shot'.

[4] Notably, they usually do not give a fine-grained account of the individual experience(s) that cause a change; and they usually cannot deal properly with repeated changes: they deal with 'one-shot' change behaviour (see footnote 3), whereas a proper change model must cope with multiple revision steps (just as a proper belief revision or learning theory must cope with multiple revisions).

[5] Out of lack of intuition one might select a very general model (with many degrees of freedom), but this may be undesirable both on theoretical (simplicity) grounds and on statistical (testability and estimatability) grounds. By contrast, selecting a highly specific model (with

I finally mention that the present analysis intends to model change *without a decay* in time of the effect of past experiences and of the individual's initial state and constitution. (But the framework is open to adding decay in future research.)

2 Change models

I define a *(change) model* as any tuple $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ consisting of:

- a set $\mathbf{S} \subseteq \mathbf{R}$ (of individual *states*), a closed interval of finite positive length,
- an arbitrary set $\mathbf{E}$ (of *experiences*) with $\mathbf{E} \neq \emptyset$,
- an arbitrary set $\mathbf{C}$ (of individual *constitutions*) with $\mathbf{C} \neq \emptyset$,
- a surjective function $(\bar{\cdot}) : \mathbf{C} \to \mathbf{S}$ (the *state projection*),
- a function $(.|.) : \mathbf{C} \times \mathbf{E} \to \mathbf{C}$ (the *revision rule*),

such that two conventions defined below (namely (1) and (2)) are respected. I now explain each component of a change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$. First though, I stress that this notion of a change model is entirely general and flexible: it abstracts from the particular characteristic of interest and might thus be applied to, for instance, (the change of) taste for music, compassion, risk aversion, drug addiction, need for recognition, impatience (discounting rate), personal or capital, and so on. Needless to say, the plausibility of a given change model or condition on change models may depend on the particular characteristic, context and individual it is applied to.

The interval $\mathbf{S}$ contains the possible states s in which the individual may be relative to the characteristic, i.e. the possible addiction levels, sympathy levels, and so on.[6]

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 structure of experiences is entirely general: they might be numbers, vectors, functions, elements of a metric space, or whatever the modeller wishes.

The set $\mathbf{C}$ contains all constitutions in which the individual may be. Constitutions are, like experiences, objects of any kind (e.g. vectors). Although one is ultimately interested in the individual's state, not in his entire constitution, constitutions will play a key role: any change in state will be explained by an underlying change in constitution. Constitutions must thus (indirectly)

---

few degrees of freedom) may be ad hoc as long as we do not have clear theoretic reasons for this selection.

[6]Under a generalised notion of a change model (left to future research), the set of states $\mathbf{S}$ might be a subset of some Euclidean space $\mathbf{R}^k$ (to model change in $k \geq 1$ characteristics), or, even more generally, *some* set endowed with *some* appropriate structure (presumably at least a topological structure so that one can formulate notions such as attraction to a state).

encode not just the current state but also the way the state would react to potential future experiences. Does this imply that constitutions must be highly complex objects, as complex as a genetic code, and too complex for practice? This question will be of central interest to us (and I hope to bring some positive news).

The state projection $(\bar{\cdot}) : \mathbf{C} \to \mathbf{S}$ assigns to each constitution $c \in \mathbf{C}$ the state $\bar{c}$ that the individual has in constitution c. So this function extracts the ultimately relevant information from constitutions. (The surjectivity assumption is there to ensure that $\mathbf{S}$ contains no impossible states.)

The revision rule $(.|.) : \mathbf{C} \times \mathbf{E} \to \mathbf{C}$, finally, maps each constitution-experience pair (c, e) to a new constitution $c|e$, read 'c after e' and representing the individual's new constitution after experiencing e in constitution c. Of ultimate interest to us is the change of *state* from $\bar{c}$ to $\overline{c|e}$.

I call the change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ a *submodel* of another one $(\hat{\mathbf{S}}, \hat{\mathbf{E}}, \hat{\mathbf{C}}, (\hat{\bar{\cdot}}), (.\hat{|}.))$ (and the latter a *supermodel* or *extension* of the former) if $\mathbf{S} \subseteq \hat{\mathbf{S}}$, $\mathbf{E} \subseteq \hat{\mathbf{E}}$, $\mathbf{C} \subseteq \hat{\mathbf{C}}$, $(\bar{\cdot}) = (\hat{\bar{\cdot}})|_{\mathbf{C}}$ and $(.|.) = (.\hat{|}.)\big|_{\mathbf{C} \times \mathbf{E}}$.

Remark. All five components of my definition are indispensable. Attempts to model change without involving constitutions run into problems (and this might explain the neglect of change in theoretical economics). Here are two such attempts that fail:

- Trying to define change directly on the level of states – namely as a mapping $\mathbf{S} \times \mathbf{E} \to \mathbf{S}$ from state-experience pairs (s, e) to new states, without involving constitutions – runs into the problem that the present state s does in practice not encode enough information to infer the post-experience state, as concrete examples make clear.[7]

- In response, one might be tempted to blow up the notion of a state such that it contains the missing information on how to update. But such 'thick states' simply collapse into what I call constitutions, and one would have to introduce a function that extracts from the 'thick state' the parameter of economic interest (e.g. the altruism level); this function corresponds to a state projection $(\bar{\cdot})$, and so we are back to a change model in the original sense up to relabelling.

Notation. I often drop brackets when it is clear how to place them; e.g. $c|e|e'$ stands for $(c|e)|e'$, and $c|e_1 \cdots |e_n$ for $(\cdots (c|e_1) \cdots)|e_n$ (interpreted as c if $n = 0$).

---

[7] Suppose states are peacefulness levels. Let an individual currently be in a peaceful state s_{old}, and then experience a war. In scenario 1, the peacefulness is not deeply entrenched in the individual, and the war experience puts him into an aggressive state $s_{\mathrm{new},1}$. In scenario 2, the peacefulness is deeply entrenched (perhaps because of 'positive' past experiences) so that the war experience leads to a new state $s_{\mathrm{new},2}$ that is less brutal than $s_{\mathrm{new},1}$ (though perhaps more brutal than s_{old}). Obviously, the war experience has changed the same state s_{old} in two different ways, depending on additional information; the missing information is what my *constitutions* are intended to capture.

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:[8]

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

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 'loosing 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 $\overline{c} = \overline{c'}$ *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 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.[9]

3   Examples and applications

I now give three formal examples of change models, followed by four concrete applications.

Example 1: the linear model. I start with a particularly important change model, later shown to have several salient properties. The *linear* change model (for the set of states $\mathbf{S}$) is the change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ with

- set of experiences $\mathbf{E} = \mathbf{S} \times (0, \infty)$; any 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 seeing a kid cry might have high s);
- set of constitutions $\mathbf{C} = \mathbf{S} \times [0, \infty)$; any 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;
- state projection given by $\overline{s_x} = s$ (so the 's' in a constitution s_x stands indeed for the current state);
- revision rule given by $s_x | \tilde{s}_{\tilde{x}} = \left(\frac{x}{x+\tilde{x}} s + \frac{\tilde{x}}{x+\tilde{x}} \tilde{s} \right)_{x+\tilde{x}}$.

---

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

[9] If $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ violates the conventions, then, as essential identity defines an equivalence on $\mathbf{E}$ (resp. $\mathbf{C}$), one can replace $\mathbf{E}$ (resp. $\mathbf{C}$) by a subset $\hat{\mathbf{E}}$ (resp. $\hat{\mathbf{C}}$) containing exactly one member of each equivalence class, $(\bar{\cdot})$ by its restriction to $\hat{\mathbf{C}}$, and $(.|.)$ by its restriction to $\hat{\mathbf{C}} \times \hat{\mathbf{E}}$.

In the linear model, 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 the weight of s resp. $\tilde{s}$ proportional to strength of constitution resp. experience. The new constitution has strength $x + \tilde{x}$. So, the stronger the old constitution and the experience, the stronger the new constitution, i.e. 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 a whole 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. Secondly, as a more abstract (measure theoretic) example, consider for a given set of states $\mathbf{S}$ the change model $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ defined by 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 $f : \mathbf{S} \to [0, \infty)$; $f(s)$ represents how much the individual currently 'tends' to state s;
- 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;
- state projection given by $\bar{c} = Median(c)$;[10] the individual's state $\bar{c}$ represents a 'summary' or 'compromise' of all tendencies in his current constitution c;[11]
- revision rule given by $c|e = c + e$.

In this change model, the individual's post-experience constitution is the sum of his old constitution c and the experience e; so that his state changes from $Median(c)$ to $Median(c + e)$.

Example 3: trivial models. Finally, as a somewhat extreme and perhaps degenerate case, 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 a single experience, i.e. that $\#\mathbf{E} = 1$, and by (2) that constitutions are isomorphic to states, i.e. that one may assume w.l.o.g. that $\mathbf{C} = \mathbf{S}$ with state projection given by $\bar{c} = c$. The model represents an individual who (at least in his environment) can have only one experience, an overwhelming one that fully overrides his initial constitution in favour of constitution c^*.

---

[10] 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).

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

I now briefly sketch potential applications.

Application 1: 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 certain goods (e.g. wine or classical music) depends on past consumption; let this be so for good 1. Specifically, let the individual's state s (in the sense of change models) be a measure of taste for good 1. If in a period t the agent consumes bundle b at 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 a coefficient attached to good 1.[12]

Becker's insight 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 – is very important, but his theory gives no clear answers to our question of *how* taste changes, i.e. on how to specify $s_{b_1,...,b_{t-1}}$ as a function of $b_1, ..., b_{t-1}$. Answers matter notably for the intertemporal consumer problem of maximising intertemporal utility

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

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

If for example one follows the linear change model, taste changes as follows. The experience of consuming any bundle b is then represented as a strength-indexed state $s(b)_{x(b)}$ (i.e. consumption of b attracts taste to $s(b)$ with strength $x(b)$). Let $\hat{s}_{\hat{x}}$ be the agent's initial constitution (i.e., initially, taste is $\hat{s}$, entrenched to degree $\hat{x}$). Applying (3), the past consumption of $b_1, ..., b_{t-1}$ leads to the new period-t taste given by:

$$s_{b_1,...,b_{t-1}} = \overline{\hat{s}_{\hat{x}}|s(b_1)_{x(b_1)} \cdots |s(b_{t-1})_{x(b_{t-1})}} = \frac{\hat{x}\hat{s} + x(b_1)s(b_1) + ... + x(b_{t-1})s(b_{t-1})}{\hat{x} + x(b_1) + ... + x(b_{t-1})}.$$

If by contrast one follows Example 2's change model, taste (hence optimal consumption paths) change rather differently. The experience of consuming any bundle b is then represented as a measure $\mu(b)$ on **S**, and consumption stream $b_1, ..., b_{t-1}$ changes the agent's initial constitution $\hat{\mu}$ (another measure on **S**) to the the period-t constitution $\hat{\mu}|\mu(b_1) \cdots |\mu(b_1) = \hat{\mu} + \mu(b_1) + ... + \mu(b_{t-1})$, whose median defines the new taste $s_{b_1,...,b_{t-1}}$.

---

[12]Using the CES utility function, $u(b; s) = ((s_1 b^1)^\rho + ... + (s_K b^K)^\rho)^{1/\rho}$ for parameters $s_1, ..., s_K, \rho > 0$, where $s_1 = s$ (taste for good 1) changes over time but $s_2, ..., s_K$ (testes for goods $2, ..., K$) and ρ stay fixed.

9

Application 2: state as level of sympathy for another player. Often observed cooperative behaviour in repeated human interactions with a prisoners'-dilemma-type monetary payoff structure is arguably best explained neither by irrational selfish agents, nor by rational agents with a stable level of sympathy (other-regardingness), but by rational agents with changing sympathy levels. Indeed, sympathy for others plausibly changes with their behaviour. But this approach requires modelling change in sympathy level. As an illustration, consider a dynamic game with perfect information, two players 1,2, 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 two actions C ('cooperate') and D ('defect'). A stage-t outcome $(B^1, B^2) \in \{C, D\}^2$ leads to monetary transfers

	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')

of the structure of a prisoners' dilemma (Table 1, left); let $v^1_{B^1B^2}$ $(v^2_{B^1B^2})$ denote the transfer, or *material* payoff, received by player 1 (2). By contrast, a player i's intra-period *utility* from (B^1, B^2) also reflect 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 $s = 0$ $(s = 1)$ represents full selfishness (sympathy).[13] Specifically, let i's intra-period utility be $u^i(B^1, B^2; s) := v^i_{B^1B^2} + s v^j_{B^1B^2}$, the sum of i's own transfer and j's transfer weighted by current sympathy.[14] Player i's intertemporal utility of a path $h \equiv (B^1_t, B^2_t)_{t=0,...,T} \in (\{C, D\}^2)^{T+1}$ is the sum of his intra-period utilities:

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

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

Once again, the core question is: how should $s^i(h_t)$ be specified, i.e. how do past experiences change the individual's state? Suppose the linear change model is used. The experience of cooperation (defect) by the other player is then represented by a strength-indexed state $s^C_{x^C}$ $(s^D_{x^D})$; so C and D are (as experiences) identified with $s^C_{x^C}$ and $s^D_{x^D}$, respectively. Let $\hat{s}_{\hat{x}}$ be each player's initial constitution, also a strength-indexed state. Two implicit assumptions (that could be dropped) are: there is symmetry between the players (i.e. the parameters $s^C, x^C, s^D, x^D, \hat{s}, \hat{x}$ are not player-indexed), and a player's *own* actions do not affect him (neglecting phenomena like habit-formation). A player's

---

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

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

experiences are thus the other player's moves. By (3), a player i's stage-t state (sympathy level) is given by

$$s^i(h_t) = \overline{\hat{s}_{\hat{x}}|B_0^j \cdots |B_{t-1}^j} = \frac{\hat{x}\hat{s} + n_t x^C s^C + (t - n_t)x^D s^D}{\hat{x} + n_t x^C + (t - n_t)x^D},$$

with n_t denoting the number of times cooperation C ($\equiv s_{x^C}^C$) occurs among $B_0^j, ..., B_{t-1}^j$. Note that we have now fully specified a dynamic game with endogenously changing mutual sympathy. For many reasonable parameter combinations[15], 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[16] in order to win the sympathy of the other player (although cooperation gives him less *current* utility), and later players cooperate because they like each other (with cooperation now dominant in the current constituent game). This contrasts with the 'always defect' outcome in a classic finitely repeated prisoners' dilemma.

Application 3: asymmetric information on 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 preference-relevant in that the path of states the characteristic undergoes enters the player's (or perhaps other players') utility function. Players with changing sympathy levels (Application 2) are just one example. It is often realistic to assume incomplete information about (i) an instable player's initial state or constitution (how initally wine-addicted is he? or how sympathetic?), and/or about (ii) the effect of future moves, i.e. moves of himself (how does his wine-drinking affect his wine-addiction?) or of others (how do my kind actions affect his sympathy for me?) or of nature. More formally, and using here the linear change model, there may be incomplete information about (i) an instable player's initial constitution $\hat{s}_{\hat{x}}$ 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 an instable player), but also (or only) the instable player himself: sometimes we are the worst judges of ourselves.

The relevance of such games is obvious, and our question becomes even more pressing: which change model should such games employ?

Application 4: explaining dynamic inconsistency by change in characteristics. Dynamic inconsistency is change in *fundamental* preference, i.e. in preference over maximally described outcomes (as opposed to Applications 1-3, to Becker's theory, and to informational preference change[17]). Models of

---

[15]Reasonably, $s^D < \hat{s} < s^C$, i.e. a defect-experience reduces sympathy, 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.

[16]The number of early stages is zero if initial sympathy $\hat{c}$ already exceeds $1/2$.

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

dynamically inconsistent agents often suffer from underdetermination and an embarrassment of riches. 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. 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 changes in preference by changes in a single characteristic, which (in our change model) are in turn explained by experiences $e \in \mathbf{E}$ (such as drug consumption, or (un)friendly actions of others, or internal experiences like Alzheimer or puberty). The question is once again: what change model should be used?

4   Two postulates on change and a consequence

Hereafter, let $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ be an arbitrary change model. 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 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 surprising because none of my postulates is 'about' the order of experience.

This 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. The co-existence of two readings – direct or by contraposition – pertains to many results, including Aumann's on agreeing to disagree.[18]

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 $\overline{c}$. An experience is* attracting *if it attracts to a state.*

---

[18]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).
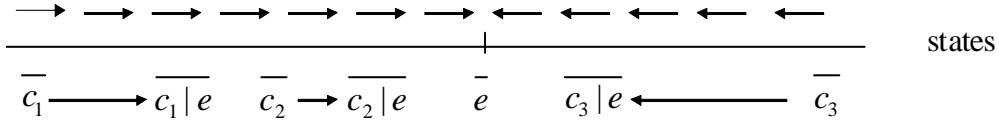
Figure 1: Experience e attracts constitutions c_1, c_2, c_3.

Two facts are worth recording (the proofs are obvious).

- If an experience e attracts to the state $\overline{c}$ of a constitution c, then the new constitution $c|e$ still has this state, i.e. $\overline{c|e} = \overline{c}$.
- Each experience e attracts to at most one state, which (if existent) is denoted $\overline{e}$ 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 $\overline{e}$.

Attraction holds in Examples 1-3, with the attractor of experiences given by $\overline{s_x} = s$ in Example 1, by $\overline{e} = Median(e)$ in Example 2, and by $\overline{e} = \overline{c^*}$ in Example 3. Attraction is a plausible (but 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, denoting for any constitution c by c_n the constitution $c|e\cdots|e$ resulting from 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 holds in Examples 1-3,[19] is, like Attraction, a 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 under highly extreme and artificial circumstances: those 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 100st experiences being any different or

---

[19]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, $\overline{c'_n} \to \tilde{s}$. And I_2 holds in Example 2 because $\overline{c_n} = Median(c + ne) \to Median(e)$ and $\overline{c_n|e'} = Median(c + ne + c') \to Median(e)$.

13

weaker than the first. Intuitively, as total experience grows, the past (in I_1) or future (in I_2) *in comparison* becomes smaller and smaller, and ultimately negligible. Part of the condition's plausibility also stems from an explicit limit in focus of this paper, which intends to model situations *without decay*: an experience e does not gradually lose its power, is not 'forgotten', as time progresses and further (similar or different) experiences are made.[20] Putting decay into the model (a challenge for future research) could be accomplished in various ways, *some* of which would involve weakening or dropping Indoctrination.[21]

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$_1$ holds in Examples 1-3, with composition given by $s_x \tilde{s}_{\tilde{x}} = \left(\frac{sx + \tilde{s}\tilde{x}}{x + \tilde{x}} \right)_{x + \tilde{x}}$ in Example 1, by $e\tilde{e} = e + \tilde{e}$ in Example 2, and by $ee = e$ in Example 3. A model that violates R$_1$ can always be enriched into one satisfying R$_1$ by simply 'closing **E** under composition'.[22]

Richness conditions are sometimes misinterpreted; so some brief remarks 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.[23] This paper uses certain conditions of richness in experiences or constitutions. If in a concrete application the

---

[20]So, in Indoctrination, each new occurrence of e is intuitively 'added' to the earlier ones, without 'replacing' or 'diminishing' them.

[21]One such way (in fact, one that 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.

[22]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, i.e. by the prescription $(c, \hat{e}) \mapsto \hat{e}(c)$.

[23]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).

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 of the real environment so as to also include what *would* happen in hypothetical cases.[24]

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

This order-invariance finding is easily checked on the models of Example 1 (where $\overline{s_x \tilde{s}_{\tilde{x}}} = \overline{\tilde{s}_{\tilde{x}} s_x} = \frac{xs + \tilde{s}\tilde{x}}{x + \tilde{x}}$ for all experiences $s_x, \tilde{s}_{\tilde{x}}$), Example 2 (where $\overline{ee'} = \overline{e'e} = Median(e + e')$ for all experiences e, e'), and Example 3. This order-invariance property holds not just in change models as simple as Examples 1-3, but surprisingly in *all* change models satisfying Attraction, Indoctrination and Richness$_1$. Full order-insensitivity, i.e. $ee' = e'e$ rather than just $\overline{ee'} = \overline{e'e}$ for all experiences e, e', is obtained in Theorem 2 below.

The rest of the section consists in proving Theorem 1.

Lemma 1 *Suppose R$_1$. Composition $\circ$ is an associative operation on* $\mathbf{E}$ *(hence* $(\mathbf{E}, \circ)$ *is a semigroup).*

So I may without ambiguity drop brackets: $ee'e''$ stands for either $e(e'e'')$ or $(ee')e''$, and e^n for the n-fold self-composition $e \cdots e$ $(n \geq 1)$.

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 R$_1$ repeatedly,

$$
\begin{aligned}
c|e(e'e'') &= c|e|e'e'' = c|e|e'|e'', \\
c|(ee')e'' &= c|ee'|e'' = c|e|e'|e''. \quad \blacksquare
\end{aligned}
$$

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 $\overline{e} \neq \overline{e'}$.*

For instance, 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 $\overline{e} = \overline{e'}$. Let c be a constitution with this state. Applying twice A and then R$_1$, we have $\overline{c} = \overline{c|e} = \overline{c|e|e'} = \overline{c|ee'}$. So $\overline{c} = \overline{c|ee'}$. Hence $\overline{ee'} = \overline{c}$ by A, i.e. $\overline{ee'} = \overline{e} = \overline{e'}$, as desired.

---

[24] 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.

2. Now suppose $\overline{e} < \overline{e'}$. Let c be a constitution of state $\overline{c} = \overline{ee'}$. I have to show that $\overline{c} \in (\overline{e}, \overline{e'}]$, and do so in two claims.

Claim 1. $\overline{c} \leq \overline{e'}$.

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

Claim 2. $\overline{c} > \overline{e}$.

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

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

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

(a) $\overline{c|e^n} \to \overline{e}$ *as* $n \to \infty$ *if I_1 holds;*
(b) $\overline{c|e^n e'} \to \overline{e}$ *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 $\overline{c'} = \overline{e}$. 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 $\overline{e}$, we have $\overline{c|e^n} \to \overline{e}$ 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) $\overline{c|e^n} \to \overline{e}$, we have $\overline{c|e^n e'} \to \overline{e}$ as $n \to \infty$. ■

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

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

First, if $\overline{e} = \overline{e'}$ then by Lemma 2 $\overline{e'e^n} = \overline{e} \to \overline{e}$ as $n \to \infty$.

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

By an analogous argument, if $\overline{e'} > \overline{e}$ then $\overline{e'e^n} \to \overline{e}$.

(b) Assume I.

First, the case that $\overline{e} = \overline{e'}$ can be treated like in part (a).

Second, suppose $\overline{e'} < \overline{e}$. Like in (a), it can be seen that $(\overline{e^n e'})_{n \geq 0}$ is an increasing sequence converging to some state $\overline{\overline{e}}$. For a contradiction, assume $s < \overline{\overline{e}}$. Letting c be a constitution of state $\overline{c} = s$, we have $\overline{c|e^n e'} \to \overline{\overline{e}}$ by part (b) of Lemma 3. So there is an $n \geq 0$ such that $[c|e^n e'] \geq \overline{c}$. But $\overline{c|e^n e'} < \overline{c}$ by $\overline{e^n e'} < \overline{c}$ and A, a contradiction.

Third, if $\overline{e'} > \overline{e}$ then for analogous reasons $\overline{e^n e'} \to \overline{\overline{e}}$.

(c) Assume I. It suffices to show that (i) $\overline{e' e^{2n} e''} \to \overline{\overline{e}}$ as $n \to \infty$ and (ii) $\overline{e' e^{2n+1} e''} \to \overline{\overline{e}}$ 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 $\overline{\overline{e}}$, so does $\overline{e' e^{2n} e''}$. ∎

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, $\overline{ee'} = \overline{e'e}$. ∎

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 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 surprisingly 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: that is, either as providing support for change models that treat experience as commutative and render experiences and constitutions isomorphic to state-strength pairs, *or* as informing us that any notion of change without

these properties can be implemented only by sacrificing some simplicity, i.e. by violating either Attraction or Indoctrination or Attraction-Consistency.

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

Attraction-Consistency (AC). This condition has two parts.
(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. $s \leq \overline{c|e} < \overline{c|e'}$ or $s \geq \overline{c|e} > \overline{c|e'}$) then *each* constitution c with $s \neq \overline{c|e'}$ is so.
(AC_2) For all constitutions c, c' of same state, if *some* experience e attracting to a state s more attracts c than c' (i.e. $s \leq \overline{c|e} < \overline{c'|e}$ or $s \geq \overline{c|e} > \overline{c'|e}$) then *each* experience e attracting to a state $s \neq \overline{c'|e}$ does so.

AC holds in the linear model and the trivial model, but not in Example 2's non-parametric model, as the reader might check. AC should be considered, more than A and I, as a genuine restriction of generality, not as an inherent property of change per se. AC is natural for applications in which two experiences (and also two constitutions) can always be meaningfully compared in terms of their strength, but less plausible in those more complex applications in which strength comparisons can be ambiguous.[25] In fact, there even is an equivalence between AC and the possibility to linearly order experiences and also constitutions in terms of strength, as can be deduced from Lemma 11 below.

Simplicity and parsimony considerations might speak in favour of adopting AC not just when AC is fully realistic. But my intension is less to argue for AC than to show its strong implications and, ultimately, to characterize the linear model.

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 a limiting 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 $\mathbf{S}$. In Example 3, all constitutions are weak.

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 $e_1, ..., e_n$ ($n \geq 1$).

Examples 1-3 obviously satisfy R_2. What makes R_2 intuitively fairly plausible is that, starting from a weak constitution (provided there is at least one), it should be possible to reach any non-weak constitution (through appropriate

---

[25]One might indeed imagine experiences e and e' none of which is unambiguously stronger than the other in the sense that e impresses the individual more or less than e' depending on the individual's current constitution. If such complicated situations are to be captured, AC should be given up.

experiences) because weakness 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* $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ *satisfies Attraction, Indoctrination, Attraction-Consistency and Richness$_{1,2}$, then* $ee' = e'e$ *for all experiences* $e, e' \in \mathbf{E}$.

That is, $(\mathbf{E}, \circ)$ is an Abelian (commutative) semigroup. To prepare 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* $\overline{c|e'}$*;*

(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 $\mathbf{E}_s$ (or $\mathbf{C}_s$) simply given by $s_x \geq s_{\tilde{x}} \Leftrightarrow x \geq \tilde{x}$, in line with our earlier interpretation of 'x' in 's_x' as strength of experience (or 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.[26]

Endowing $\mathbf{E}_s$ and $\mathbf{C}_s$ with their strength relations yields structures $(\mathbf{E}_s, \geq)$ and $(\mathbf{C}_s, \geq)$; further endowing $\mathbf{E}_s$ with composition $\circ$ (under R_1) yields a structure $(\mathbf{E}_s, \geq, \circ)$ (an ordered semi-group, as we will see). Isomorphisms between structures (i.e. between sets 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. (Trivial models were defined in Example 3.)

---

[26] 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 $\mathbf{C}_s$.

Theorem 3 *If a non-trivial change model* $(\mathbf{S}, \mathbf{E}, \mathbf{C}, (\bar{\cdot}), (.|.))$ *satisfies Attraction, Indoctrination, Attraction-Consistency and Richness$_{1,2}$, then for every state 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 $(0, \infty) \cap \mathbf{Q}$ 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, and this submodel satisfies all conditions of Theorem 3, namely with $X_s = X$ for all $s \in \mathbf{S}$. In other examples, the set X_s varies across states 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.

I now prove Theorems 2 and 3, starting with several lemmas.

Lemma 5 *Assume R_1.*

(a) *Constitutions c, c' are identical if $\overline{c} = \overline{c'}$ 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 $\overline{c} = \overline{c'}$ and $\overline{c|.} = \overline{c'|.}$. By (2) and as $\overline{c} = \overline{c'}$, 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 $\overline{c|e\tilde{e}} = \overline{c|e'e}$ 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). ∎

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 $\overline{e}$ and $\overline{e'}$.*

20

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 $\overline{e}$ and $\overline{e'}$. Also by Lemma 2, $\overline{ee'} \neq \overline{e}$ and $\overline{e'e} \neq \overline{e'}$, the latter implying $\overline{ee'} \neq \overline{e'}$ by Theorem 1. So $\overline{ee'}$ is strictly between $\overline{e}$ and $\overline{e'}$. ∎

Adding R_1 to R_2, all constitutions c are reachable from the *same* weak constitution c_w, and this through a *single* experience:

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 $\overline{c_w|e_ce} = \overline{c'_w|e_ce}$ by R_1, hence to $\overline{e_ce} = \overline{e_ce}$ 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 4. 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. ∎

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 that $\overline{\hat{e}ee'} = \overline{\hat{e}e'e}$ for all $\hat{e} \in \mathbf{E}$. So 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 $\hat{e}ee' = \hat{e}e'e$.

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

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.[27] 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 theorem, a number of lemmas must be shown.

Lemma 10 *Assume A and AC.[28] 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.

(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 $R_{1,2}$.*[29] *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 $\hat{e}e = \hat{e}e'$ 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. $\overline{c_w|\hat{e}|e'} = \overline{c_w|\hat{e}|e'}$. So by R_1 $\overline{c_w|\hat{e}e'} = \overline{c_w|\hat{e}e'}$, and hence by c_w's weakness $\hat{e}e' = \hat{e}e'$, as desired.

(b) Consider constitutions $c, c' \in \mathbf{C}_s$ such that $c \equiv c'$. Then, for all experiences 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 $e \in \mathbf{E}$. Hence $c = c'$ by Lemma 5. ∎

Note the large 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 $\mathbf{R}$, the lexicographically ordered set $\mathbf{R}^2$ 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).

1. 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.

2. 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. ∎

Lemma 13 *Assume A, I, AC and $R_{1,2}$.*[30] *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 $\{c \in \mathbf{C} : c \text{ is not weak}\}$ and*
- *for each state s an (order-)isomorphism between $(\mathbf{E}_s, \geq)$ and $(\{c \in \mathbf{C}_s : c \text{ is not weak}\}, \geq)$.*

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.

Consider 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 $\overline{e_1 e} \neq \overline{e_2 e}$. So, by the weakness of c_w, $\overline{c_w|e_1 e} \neq \overline{c_w|e_2 e}$. Hence, $c_w|e_1 e \neq c_w|e_2 e$, 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 $\{c \in \mathbf{C}_s : c \text{ is not weak}\}$. 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, $\mathbf{E}_s$ is closed under $\circ$, i.e. $\circ$ 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 have to 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 $\overline{ee'\hat{e}}$. But this implies, again by Lemma 12, that $\dot{e} > e$. Contradiction. Q.e.d.

Claim 2. $(\mathbf{E}_s, \geq, \circ)$ is positively ordered.

---

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

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 $\overline{\ddot{e}\hat{e}}$ 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 $\overline{\ddot{e}\hat{e}} \neq s$ by Lemma 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. $\circ$ 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

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

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 $\overline{\ddot{e}} \neq \overline{e}$. 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. Then, by Lemma 8, $\mathbf{E} = \{e\}$. Hence, by non-triviality, there is a $c \in \mathbf{C}$ such that (*) $\overline{c|e} \neq \overline{e}$. In particular, 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 $\mathbf{E} = \{e\}$ we have $e_c = e$, and hence $\overline{c} = \overline{c_w|e} = \overline{e}$. So $\overline{c|e} = \overline{e}$, a contradiction by (*). 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.[31] 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 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, $e^n > e'^{n+1}$. ∎

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-Nakadathe 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.

---

[31] 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 $3 = 2 + x$.

To show the second bullet point, write $\mathbf{C}_s^* := \{c \in \mathbf{C}_s : c \text{ is not week}\}$. 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, $(\mathbf{C}_s, \geq) = (\mathbf{C}_s^* \cup \{c_w\}, \geq) \equiv (X_s \cup \{0\}, \geq)$. $\blacksquare$

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 three more 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 $e \in \mathbf{E}$).[32] Isomorphic models are perfectly equivalent (but perhaps not equally natural or convenient).[33]

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

Richness$_3$ (R$_3$) For each constitution c there are experiences e, e' such that $\overline{c|e} \leq \overline{c} \leq \overline{c|e'}$.

Richness$_4$ (R$_4$) No non-weak constitution c is weaker than all other non-weak constitutions of the same state, i.e. satisfies $c < c'$ for all other non-weak constitutions c' of state $\overline{c'} = \overline{c}$.

Intuitively, R$_3$ requires that change is never 'one way only', i.e. that any state could, depending on the experience, either (weakly) fall or (weakly) increase – a mild form of richness in experiences that holds whenever some experience in $\mathbf{E}$ attracts to the maximal state and some to the minimal state. R$_4$ holds in

---

[32]For instance, the linear model with state set $\mathbf{S} = [0,1]$ is isomorphic to that with state set $\hat{\mathbf{S}} = [0,2]$: 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 scale obtained by squaring each strength level.

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

particular if for any (non-weak) constitution there is a weaker one (without the strength relation $\geq$ being assumed to be linear).

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 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. The set of all state paths, $\mathbf{S}^{\cup_{n=0}^{\infty} \mathbf{E}^n}$, usually contains many state paths that are impossible, i.e. are not the state path of any constitution c in $\mathbf{C}$. A state path $(s_{\mathbf{e}})_{\mathbf{e} \in \cup_{n=0}^{\infty} \mathbf{E}^n}$ is *constant* if its states $s_{\mathbf{e}}$ are all the same.

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.[34] 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 $\mathbf{S}^{\cup_{n=0}^{\infty} \mathbf{E}^n}$.[35]

We are ready for the characterization result.

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

The proof 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.[36]

---

[34]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 $n \to \infty$.

[35]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 $\mathbf{e} \in \cup_{n=0}^{\infty} \mathbf{E}^n$).

[36]An ordered semigroup $(X, \geq, \circ)$ is *dense* if $\geq$ is dense (i.e. for all $x, y \in X$ with $x > y$

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 $S \in \{\mathbf{R}, [0, \infty), (0, \infty), [1, \infty), (1, \infty)\}$ 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 non-trivial. 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 $(e_k)_{k=1,2,\ldots}$ 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 $\{e_k : k = 1, 2, \ldots\}$ 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 $\mathbf{R}$): 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 $(c_w|e_k)_{k=1,2,\ldots}$, i.e. the sequence of state paths whose k's component is $\overline{(c_w|e_k|e_1 \cdots |e_n)}_{(e_1,\ldots,e_n) \in \cup_{n=0}^\infty \mathbf{E}^n} = (\overline{e_k e_1 \cdots e_n})_{(e_1,\ldots,e_n) \in \cup_{n=0}^\infty \mathbf{E}^n}$, 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 is analogous if $\overline{e} < s$. 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, $\overline{c_*|e} < \overline{e}$. So $\overline{c_*|e} \neq \overline{e}$. 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 $\{e_k : k = 1, 2, \ldots\}$ (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 $e \in \mathbf{E}$, $\overline{ee_k}$ is (by Claim 1) weakly between s and $\overline{c_*|e}$, i.e. (by Claim 2) weakly

_____

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).

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 $e \in \mathbf{E}$ such that $\overline{e_{**}e}$ is strictly between s and $\overline{e_*e}$. So there is (by Claims 1-2) a k such that $\overline{e_{**}e}$ is strictly between s and $\overline{e_ke}$; but this violates that $e_k \geq e_{**}$. $\blacksquare$

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

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 $\inf_{e' \in \mathbf{E}_s} |\overline{ee'} - \overline{e}| = 0$. Now suppose $e \notin \mathbf{E}_s$; w.l.o.g. let $\overline{e} < s$ (the proof being analogous if $\overline{e} > s$). So I have to show that $\inf_{e' \in \mathbf{E}_s} (\overline{ee'} - \overline{e}) = 0$, i.e. that $\inf_{e' \in \mathbf{E}_s} \overline{ee'} = \overline{e}$.

Claim 1. There is a $c_* \in \mathbf{C}$ such that $\inf_{e' \in \mathbf{E}_s} \overline{ee'} = \overline{c_*|e}$.

Consider a sequence $(e_k)_{k=1,2,\dots}$ in $\mathbf{E}_s$ such that $\overline{ee_k} \to \inf_{e' \in \mathbf{E}_s} \overline{ee'}$ 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 $\overline{e_ke}$ converges to $\overline{c_*|e}$ as $k \to \infty$; hence $\inf_{e' \in \mathbf{E}_s} \overline{ee'} = \overline{c_*|e}$. Q.e.d.

Claim 2. c_* is weak (hence by Claim 1 $\inf_{e' \in \mathbf{E}_s} \overline{ee'} = \overline{e}$, 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 $\inf_{e' \in \mathbf{E}_s} \overline{ee'} = \overline{c_*|e} = \overline{ee_*}$. 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. $\blacksquare$

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 $\overline{\tilde{e}} < s$ (the proof is analogous if $\overline{\tilde{e}} > 0$). Let $\epsilon := \overline{\tilde{e}e} - \overline{\tilde{e}}$ (> 0). By $\epsilon > 0$ and Lemma 21, there is an $e' \in \mathbf{E}_s$ such that $\overline{\tilde{e}e'} - \overline{\tilde{e}} \leq \epsilon/2$; 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 $(\overline{\tilde{e}e'} - \overline{\tilde{e}}) + (\overline{\tilde{e}e'e''} - \overline{\tilde{e}e'}) \leq \epsilon/2 + \epsilon/2$, i.e. $\overline{\tilde{e}e'e''} - \overline{\tilde{e}} \leq \epsilon$. So, letting $\hat{e}$ be the weakest of e' and e'', $\overline{\tilde{e}\hat{e}^2} - \overline{\tilde{e}} \leq \epsilon$, i.e. $\overline{\tilde{e}\hat{e}^2} - \overline{\tilde{e}} \leq \overline{\tilde{e}e} - \overline{\tilde{e}}$. So $\hat{e}^2 \leq e$. $\blacksquare$

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 $(\mathbf{R}, \geq)$, hence exists but needn't belong to X_s).

By Lemma 22 there exists a sequence $(e_k)_{k=1,2,...}$ 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 $(x_k)_{k=1,2,...}$ 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 ($n \in \{1, 2, ...\}$) 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 topological density in $(0, \infty)$. 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);
- $\{z \in X : z < y\} = \cup_{z \in X}\{z \in X : y < x < z\}$ 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 $|(a + b) - (a' + b')| < \epsilon$. ∎

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

Lemma 25 *Assume A, I, AC and $R_{1,2,4,5}$. 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 and $R_{1,2,4,5}$, 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 $S \in \{\mathbf{R}, [0, \infty), (0, \infty), [1, \infty), (1, \infty)\}$. But, as $e \circ e > e$ for all $e \in \mathbf{E}_s$, $(\mathbf{E}_s, \geq, \circ)$ is isomorphic to neither $(\mathbf{R}, \geq, +)$, 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 $(\mathbf{C}_s, \geq) \equiv ([0, \infty), \geq)$. $\blacksquare$

Lemma 25 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 25 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 $a > 0$, let e^a ('e raised to the power a') be the (by $(\mathbf{E}_{\overline{e}}, \geq, \circ) \equiv ((0, \infty), \geq, +)$ uniquely existing) 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 (by $(\mathbf{E}_{\overline{e}}, \geq, \circ) \equiv ((0, \infty), \geq, +)$ uniquely existing) experience $e' \in \mathbf{E}_{\overline{e}}$ given by $e'^n = e^m$ (i.e. by $\underbrace{e' \cdots e'}_{n} = \underbrace{e \cdots e}_{m}$).

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 $(0, \infty)$.[37] 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 26 *Assume A, I, AC and $R_{1,2,4,5}$. 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 $R_{1,2,4,5}$. 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

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

---

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

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 $\frac{m}{n} > a - r$. We have $\tilde{e} < e^{a-r}\dot{e}^{a-r} < e^{a-r}\dot{e}^{a-r}e^{\frac{m}{n}-(a-r)}\dot{e}^{\frac{m}{n}-(a-r)} = e^{\frac{m}{n}}\dot{e}^{\frac{m}{n}}$. ∎

Lemma 27 *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 $f : (0,\infty) \to \mathbf{R}, a \mapsto \overline{e_0 e_1^a}$. 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 26 $\overline{e_0^b e_1} = \overline{(e_0^b e_1)^{1/b}} = \overline{e_0 e_1^{1/b}} = f(\frac{1}{b})$, 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 $f(a^*) \leq s$ and $f(a^*) \geq s$.

I first show $f(a^*) \leq s$. For all $n \in \{1, 2, ...\}$, by Lemma 26 $f(\frac{na^*}{n+1}) = \overline{e_0 e_1^{na^*/(n+1)}} = \overline{e_0^{n+1} e_1^{na^*}} = \overline{e_0 (e_0 e_1^{a^*})^n}$. 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 26 $f(\frac{(n+1)a^*}{n}) = \overline{e_0 e_1^{(n+1)a^*/n}} = \overline{e_0^n e_1^{(n+1)a^*}} = \overline{e_1^{a^*} (e_0 e_1^{a^*})^n}$. 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 28 *Assume A, I, AC and $R_{1\text{-}5}$. We have $\{\overline{e} : e \in \mathbf{E}\} = \mathbf{S}$.*

Proof. Assume A, I, AC and $R_{1\text{-}5}$. Let $s \in \mathbf{S}$. I have to show that $\overline{e^*} = s$ for an $e^* \in \mathbf{E}$. Let $c \in \mathbf{C}$ be such that $\overline{c} = s$. By R_3 there are $e, e' \in \mathbf{E}$ such that $\overline{c|e} \leq \overline{c} \leq \overline{c|e'}$. So, by A, $\overline{e} \leq \overline{c} \leq \overline{e'}$. Hence by Lemma 27 there is an $e^* \in \mathbf{E}$ such that $\overline{e^*} = \overline{c} = s$. ∎

Proof of Theorem 4. First, the linear model for a state set $\mathbf{S}$ obviously satisfies all of properties A, I, AC and $R_{1\text{-}5}$; and so do its isomorphic variants, because reparametrisations 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 $R_{1\text{-}5}$, and show that it is a reparametrisation of the linear model, to be denoted $(\mathbf{S}, \hat{\mathbf{E}}, \hat{\mathbf{C}}, (\hat{\cdot}), (.\hat{|}.))$. 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 25 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 28).

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

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

34

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 27. 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 25) an $x > 0$ such that $e = [\epsilon(s_1)]^x$. So

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

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 these two observations:

- 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) mapping $\epsilon : \hat{\mathbf{E}} \to \mathbf{E}$ and the (by Lemma 13 bijective) mapping $e \mapsto \tilde{w}|e$ from $\mathbf{E}$ to $\{c \in \mathbf{C} : c \text{ is not weak}\}$.
- 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 26, 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 $(.|.)$ 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 26,

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

7 Conclusion

I hope that the axiomatic approach to modelling change is a fruitful step towards filling the wide gap between recognised importance of (endogenous or exogenous) change and lack of theoretical understanding of how to model it. Our findings can be applied in a range of ways, depending on interests and views:

- The decision- or game-theorist might *either* model a form of change that keeps fundamental preferences fixed (in line with orthodox methodology; see Applications 1-3), *or* model a form of change that induces dynamic inconsistency (see Application 4).
- He might *either* take our theorems as welcome arguments for neglecting the order of experience (Theorems 1 and 2) and perhaps for modelling change linearly (Theorems 3 and 4), *or* insist in order-relevance and non-linearity, in which case he will have to decide whether to give up Attraction, Indoctrination or Attraction-Consistency.

- The empirical researcher might take a given change model and conduct some intriguing estimation or hypothesis testing, thereby informing us about the real-life values of model parameters 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, generalise change models to multi-dimensional individual states (in order to study simultaneous change in multiple 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. My ultimate hope is that change will be taken more seriously in our modelling practice.

8 References

AUMANN, R. (1976) Agreeing to disagree, *The Annals of Statistics* 4(6): 1236-1239

ALIMOV, N. G. (1950) On ordered semigroups, *Izv. Akad. Nauk SSSR. Ser. 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, *American Economic Review* 90: 166-193

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

CLIFFORD, A. H. (1954) Naturally totally ordered commutative semigroups, *Amer. J. Math.* 76: 631-646

CLIFFORD, A. H. (1958) Totally ordered commutative semigroups, *Bull. Amer. Math. Soc.* 64(6): 305-316

DEKEL, E., J. ELY AND O. YILANKAYA (2007) Evolution of preferences, *The Review of Economic Studies* 74: 685-704

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

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

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

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

HAMMOND, P. (1976) Changing tastes and coherent dynamic choice, *Review of Economic Studies* 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

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, *American Economic Review* 89: 103-124

POLAK, R. (1976) Interdependent preferences, *American Economic Review* 66(3): 309-20

RABIN, M. (1993) Incorporating fairness into game theory and economics, *American Economic Review* 83: 1281-1302

RABIN, M. (1998) Psychology and economics, *Journal of Economic Literature* 36(1): 11-46

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

SETHI, R. AND E. SOMANATHAN (2003) Understanding reciprocity, *Journal of Economic Behavior and Organization* 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, *American Economic Review* 67: 76-90

STROTZ, R. H. (1955-56) Myopia and inconsistency in dynamic utility maximization, *Review of Economic Studies* 23(3): 165-180

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