Applied Macroeconomic Analysis with Epstein Zin Utility

Christopher Heiberger, Halvor Ruf

December 16, 2014

We gratefully acknowledge financial support from the German Research Foundation (Deutsche Forschungsgemeinschaft). This work is part of a research project within the Priority Programme 1578 "Financial Market Imperfections and Macroeconomic Performance".

This paper provides a self contained guide to the approximate solution of DSGE macroeconomies that feature Epstein and Zin (1989) (EZ) utility. We first summarize the crucial elements of the EZ representation as it is increasingly applied in the literature adding intuitive remarks and illustrative examples. Thereafter, we demonstrate within a basic representative agent economy how EZ utility naturally lends itself to dynamic programming and apply the Schmitt-Grohe and Uribe (2004) approach to find a second order perturbation. We conclude by discussing the immediate implications of employing the EZ representation in applied work by computing an actual numerical example. The paper is accompanied by a flexible `Maple–Matlab` perturbation toolbox.

Contents

List of figures

1 Introduction

To us, there are essentially three motivations for applied macroeconomists to study the Epstein and Zin (1989) (EZ) utility representation and its (incomplete) separation of the elasticity of intertemporal substitution from the standard risk aversion parameter. First, it provides the researcher with an additional degree of freedom to improve on the empirical performance of his DSGE models. Second, it is theoretically appealing to loosen those two aspects of preferences because, a priori, there does not seem to be a reason for their reciprocity as implicitly assumed in the paradigmatic framework of additively separable expected utility. Third, probably mostly for these two reasons, the applied literature is increasingly employing EZ utility.[1]

Now, although there is work on both the rationale behind the EZ representation and its approximation,[2] we find the literature to lack a unified approach which guides the reader towards the derivation of the equilibrium conditions of an EZ economy from applying the EZ utility to its primitive consumption space as it is naturally generated in basic representative agent stochastic control systems. This paper mainly intends to fill this gap and to additionally provide guidance with respect to the expected benefits for applied research.

The remainder is organized as follows. Section 2 and 3 summarize the crucial elements of the EZ representation. The material is enhanced with intuitive remarks and illustrative examples. Section 4 demonstrates the application of EZ utility to a standard representative agent decision problem as well as the application of the Schmitt-Grohe and Uribe (2004) second order perturbation approach for its approximate solution. A discussion on some immediate implications of the EZ representation for applied work is provided within a numerical exemplification. The employed perturbation routines are collected in a very flexible `Maple–Matlab` toolbox which is briefly documented in section 5.

[1]Note the introductory remarks in van Binsbergen et al. (2012) and the sources cited therein.

[2]See e.g. Backus, Routledge, and Zin (2008) or Altug and Labadie (2008) for a description of the representation or Caldara et al. (2012) on the approximation techniques.

2 Consumption space

This section introduces the key notion of temporal lotteries and their respective identification as a pair of current consumption and a probability distribution over future temporal lotteries. The inherent recursiveness of this identification will give rise to a recursive utility representation of preferences over such temporal lotteries in section 3.

2.1 Preliminary remarks

Most of the applied DSGE literature relies on a common and thus standard framework of temporal decision making under uncertainty: choice is made between probability distributions over consumption sequences, i.e. over stochastic consumption processes.

The underlying idea of the approach applied in the present paper is the introduction of additional structure to the fundamental consumption decision problem. Not only is the probability distribution of consumption sequences considered but also the time at which the uncertainty concerning future consumption is resolved. This is done via the concept of probability trees, so called *temporal lotteries*.[3] This idea and the ensuing construction of the consumption space was introduced by Kreps and Porteus (1978) and extended to an infinite horizon setting by Epstein and Zin (1989). While the former provide an axiomatization for a recursive utility representation over finite temporal lotteries, the latter prove the existence of a recursive utility function over some space of infinite horizon temporal lotteries. This section is concerned with the presentation of these ideas and is thereby intended to summarize particularly crucial results. We thereby often sacrifice the generality of the original treatment in order to keep a focus on applicability. This will provide us with the necessary basis for the application of the EZ approach to neoclassical macroeconomic analysis in the remainder of this paper.

[3] Accordingly, their atemporal analog mentioned above will sometimes be called an *atemporal lottery*.

2.2 Notation

Let X be a metric space. Denote by

$$\mathcal{B}(X) := \sigma\left(\{O \subset X |\ O \text{ open}\}\right)$$

the induced Borel σ–algebra on X and by

$$\mathcal{M}(X) := \{p : \mathcal{B}(X) \to [0,1] |\ p \text{ probability measure}\}$$

the set of Borel probability measures. Particularly, for $x \in X$ let

$$\begin{array}{rccl} \delta_x : & \mathcal{B}(X) & \to & [0,1] \\ & B & \mapsto & \delta_x(B) := \begin{cases} 1, & x \in B \\ 0, & x \notin B \end{cases}, \ \forall B \in \mathcal{B}(X). \end{array}$$

be the Dirac probability measures. Moreover, time is discrete and the planning horizon is infinite. Hence, by $t \in \mathbb{N}$ we denote a point in time or its respective period.

2.3 Temporal lotteries

Let D denote the space of temporal lotteries. Elements of D can be pictured as infinite probability trees and can thus naturally be identified with a tuple of current consumption and a probability distribution over nodes of infinite probability trees emanating next period. Accordingly, we will shortly find the space of temporal lotteries D to be homeomorphic to $\mathbb{R}_+ \times \mathcal{M}(D)$. In order to motivate the analysis of temporal lotteries, consider the following example.[4]

Example 1 Robinson Crusoe is facing the problem of choosing between two seed technologies. Both possible technologies offer an aggregate period t harvest (output) Y_t as a response to this period's aggregate input K_t according to

$$Y_t = A_t \cdot K_t, \text{ where } A_2 = \begin{cases} 2, & \text{with probability } \frac{1}{2} \\ 1, & \text{with probability } \frac{1}{2} \end{cases} \text{ and } A_t = 2\ \forall t \neq 2.$$

[4]This example is a modified version of the coin flip example originally provided by Kreps and Porteus (1978). We use our formulation instead for the sake of exposition.

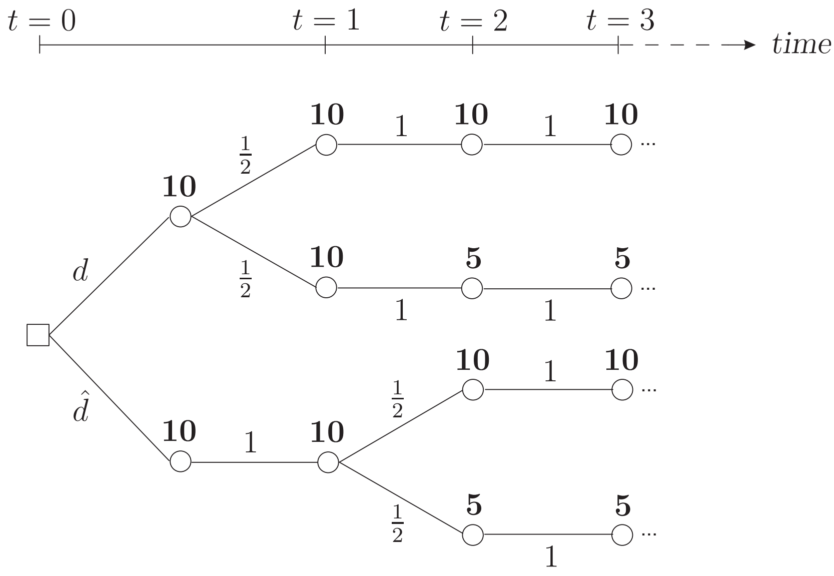

Figure 1: Temporal decision problem

The difference between these two technologies is the time at which the technologies' period 2–types are revealed.

Suppose for the first technology, its period type is revealed at the beginning of period 1 whereas for the second technology Robinson does not know about its type before period 2. Further suppose that for all periods, Robinson is rigidly bound to a somewhat primitive allocation as he distributes a constant fraction $\frac{1}{2}$ of (produced) resources to both consumption and next period's capital stock.

He now decides between these two alternative technologies according to their resulting future consumption prospects, his utility argument. We identify such a decision between actions with the decision between the probability trees induced by these actions. The decision problem is pictured in figure 1, where d and $\hat{d}$ denote the consumption probability trees that correspond to an initial endowment of 20.[5] Observe that from a period 0 point of view, both technologies result in the same *atemporal* distribution over future consumption, namely $\frac{1}{2}\delta_{(10,10,\ldots)} + \frac{1}{2}\delta_{(10,10,5,5,\ldots)}$. Hence, if we take the consumption space to be the space of atemporal lotteries over consumption streams—as it is implied by the standard model—there is no way to distinguish between

[5]In such graphs, squares denote action nodes while circles denote uncertainty nodes, cf. Raiffa (1970), p. 11.

the consequences of these alternative technologies. However evidently, the two pictured probability trees are not identical.

We conclude this introductory example with a brief discussion on its strength as a motivation for the upcoming analysis. For the case of lotteries over *income* sequences there is very little controversy about whether different temporal lotteries inducing identical atemporal lotteries ought to be modelled in a way that allows the decision maker to prefer one over the other.[6] For *consumption* sequences, however, the situation is less evident. To see this, take a second look at our example. If Robinson allocated consumption according to an optimality criterion such as the maximization of his lifetime utility, it would typically be of value to know about the future production in advance in order to improve on the allocation by situational consumption adaption.[7] However, there is no apparent planning advantage for consumption sequences, as the one just pictured. We eventually provide two remarks on that. First, even if there is no immediate reason for why it *must* be the case that a decision maker would favor prior or later resolution of consumption uncertainty, it also seems odd to insist on the view that any decision maker would *never* be able to appreciate it. Above all, introspection enhances the latter doubts. As Epstein (1992), p. 23, notes e.g. later knowledge about the future to come might very well be preferred by a person who likes to "defer resolution in order to [be able to sustain] the hope [...] for a favorable outcome for a risky prospect." Moreover, the "rationality" of nonindifference towards the timing of resolution of consumption uncertainty is nicely exemplified in Chew and Epstein (1989). Second, as it turns out, such nonindifference allows to loosen the strict entanglement of risk attitudes and intertemporal consumption substitution as implied by the standard model. Beside its theoretical appeal, this leaves macroeconomists with an additional degree of freedom in replicating empirical data. A preference for earlier or later resolution of consumption uncertainty can hence also be interpreted as a cost of the last mentioned advantages, cf. Epstein, Farhi, and Strzalecki (2014).

The construction of the consumption space carried out in Epstein and Zin (1989), p. 941-944, is mathematically involved. Since the ideas behind are nevertheless indispensable for our intended discussion of the actual appli-

[6] Thereby, preference for income sequences is understood as being induced from the primitive preference for consumption sequences as the ultimate source of utility.

[7] This point is nicely illustrated in Spence and Zeckhauser (1972).

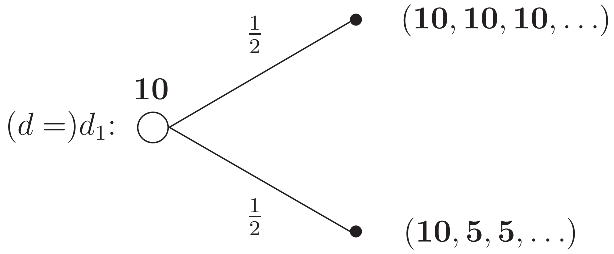

Figure 2: Temporal lottery with no uncertainty after period 1

cation of EZ utility to a basic DSGE economy, the remainder of this section summarizes their treatment in detail and complements it with some examples and additional intuition. Still, at some places more rigorous remarks supplement our summary.

2.3.1 Infinite trees of finite length

The example above shows the following. If we identify each branch of a probability tree with a consumption sequence and if—as in the case of the example probability tree d—there is no uncertainty after period 1, then no structure is lost by simply considering the "finite" probability tree d_1 that results through considering only the probability distribution of the consumption sequence starting at period 1 next to initial consumption. We thus say that the infinite tree d_1 has length 1 and generally call an infinite tree of finite length t a finite (t-stage) tree. Since we only consider infinite horizon decisions, this should not cause confusion but shorten the ensuing analysis.

In particular, figure 2 reveals that the probability tree d_1 is unambiguously represented by the tuple (c_0, m_1) consisting of current consumption $c_0 = 10$ and the probability measure $m_1 = \frac{1}{2}\delta_{(10,10,10,\ldots)} + \frac{1}{2}\delta_{(10,5,5,\ldots)} \in \mathcal{M}(\mathbb{R}_+^\infty)$ over future consumption sequences. Accordingly, we identify the infinite probability tree d with the finite tree $d_1 = (c_0, m_1) \in \mathbb{R}_+ \times \mathcal{M}(\mathbb{R}_+^\infty)$.

If we now turn to the infinite probability tree $\hat{d}$ instead (cf. figure 3), we observe that in this case there is no uncertainty after period 2. Analogously, $\hat{d}$ can be pictured as a finite two–stage probability tree $\hat{d}_2$. The one–stage tree emerging at its second node, $\hat{d}_1$, is similarly identified with the tuple $(\hat{c}_1, \hat{m}_1)$ out of period 1's consumption level $\hat{c}_1 = 10$ and the probability distribution $\hat{m}_1 = \frac{1}{2}\delta_{(10,10,\ldots)} + \frac{1}{2}\delta_{(5,5,\ldots)} \in \mathcal{M}(\mathbb{R}_+^\infty)$ over consumption as of period 2, i.e. $\hat{d}_1 = (\hat{c}_1, \hat{m}_1) \in \mathbb{R}_+ \times \mathcal{M}(\mathbb{R}_+^\infty)$. Consequently, the whole two–stage probability

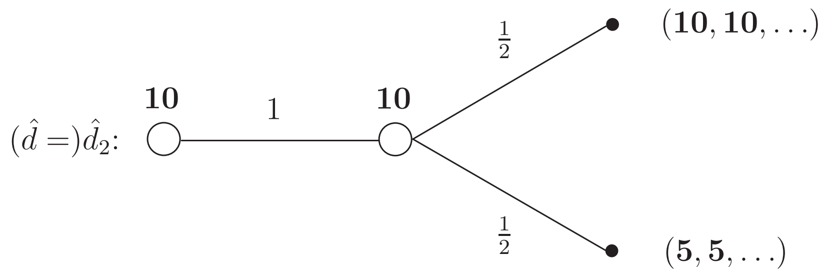

Figure 3: Temporal lottery with no uncertainty after period 2

tree $\hat{d}_2$ now comes up to a tuple of current consumption $\hat{c}_0 = 10$ and a degenerate probability distribution over nodes of one–stage trees $\hat{m}_2 = \delta_{\hat{d}_1} \in \mathcal{M}\left(\mathbb{R}_+ \times \mathcal{M}(\mathbb{R}_+^\infty)\right)$. I.e. we have $\hat{d}_2 = (\hat{c}_0, \hat{m}_2) \in \mathbb{R}_+ \times \mathcal{M}\left(\mathbb{R}_+ \times \mathcal{M}(\mathbb{R}_+^\infty)\right)$.

In general, continuing this reasoning inductively we can define finite probability trees of length t. In such trees, the way in which uncertainty resolves over time is only displayed until period $t-1$ and the only information encoded about future periods' consumption as of t is their joint probability distribution. Precisely, a probability tree of length t can be described completely by a pair of today's consumption and a probability measure over nodes of trees of remaining length $t-1$. I.e. we recursively define

$$D_0 := R_+^\infty,$$
$$D_t := \mathbb{R}_+ \times \mathcal{M}(D_{t-1}) \text{ f.a. } t \geq 1.$$

As mentioned above, we will from time to time add more rigorous remarks to round off our treatment. In this spirit, note that since R_+^∞ is a complete, separable metric space, i.e. Polish, so is $\mathcal{M}(R_+^\infty)$ with the weak topology. Therefore, $D_1 = R_+ \times \mathcal{M}(R_+^\infty)$ is also a Polish space with the product topology on it. Inductively it follows that D_t is a Polish space f.a. $t \geq 1$ if we recursively endow each $\mathcal{M}(D_{t-1})$ with the weak topology and D_t with the induced product topology. For each $t \in \mathbb{N}$ we denote by $\mathcal{B}_t := \mathcal{B}(D_t)$ the respective Borel σ–algebra.

Observe that any infinite probability tree which does not contain any uncertainty from period t on can be represented by such a t–stage probability tree without loss of information.

2.3.2 Infinite trees of arbitrary length: consistent reduction

The idea behind formally defining an arbitrary infinite probability tree d is to approximate it stepwise by the t–stage probability trees, which arise from d by reducing its temporal structure in such a way that the reduced tree has the identical distribution over consumption sequences as of period t. Since the reduced tree does not contain any information of how the uncertainty regarding these consumption sequences as of t resolves over time, considering the reduced tree is like pretending that all uncertainty about consumption will have been resolved by period t.

This way we get a sequence $(d_1, d_2, \ldots)$ of finite probability trees $d_t \in D_t$, each describing the structure of d with increasing accuracy. Since for each $t \in \mathbb{N}$, the probability tree d_t describes the temporal structure of d up to period t, the whole sequence $(d_1, d_2, \ldots)$ describes the entire structure of d. Therefore, we can identify each infinite probability tree with exactly one such sequence. Accordingly, we define the set D of all infinite probability trees as the set of all sequences

$$d = (d_1, d_2, \ldots), d_t \in D_t \text{ f.a. } t \in \mathbb{N},$$

which are consistent in the following sense. For $d_t \in D_t$ and $d_{t+1} \in D_{t+1}$ to be consistent it must hold that up until period t they both obey the same structure and that d_t can be imagined as d_{t+1} folded back one period. To put it another way, d_t and d_{t+1} are consistent if d_t results from d_{t+1} by neglecting its temporal structure already in period t and merely considering the joint distribution of consumption to come. The process of consistently reducing probability trees is first illustrated in the following example and then outlined generally.

Example 2 Consider a two-stage probability tree. As depicted in figure 4, consistent reduction of such a tree demands the computation of the induced distribution of consumption to come as of period 1. Note first that e.g. the two-stage tree pictured in figure 5 is trivially consistently reduced to the same D_1 tree as the two-stage tree of figure 4. Hence, consistent reduction cannot be an injection. In particular, for all trees that share this consistent reduction, the induced probability measure must e.g. for $B = \{(c_1^1, c_2^1, 1, 1, \ldots), (c_1^3, c_2^4, 1, 1, \ldots)\} \in \mathcal{B}\left(\mathbb{R}_+^\infty\right)$ yield

$$P(B) = p_0^1(1 - p_1^1) + (1 - p_0^1 - p_0^2)p_1^2.$$

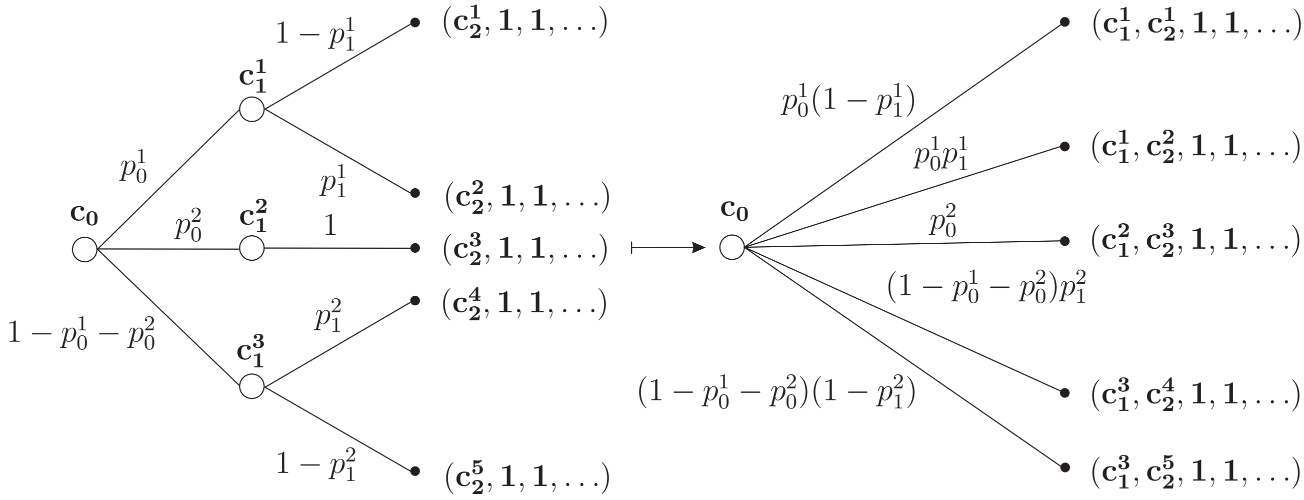

Figure 4: Consistent reduction (example)

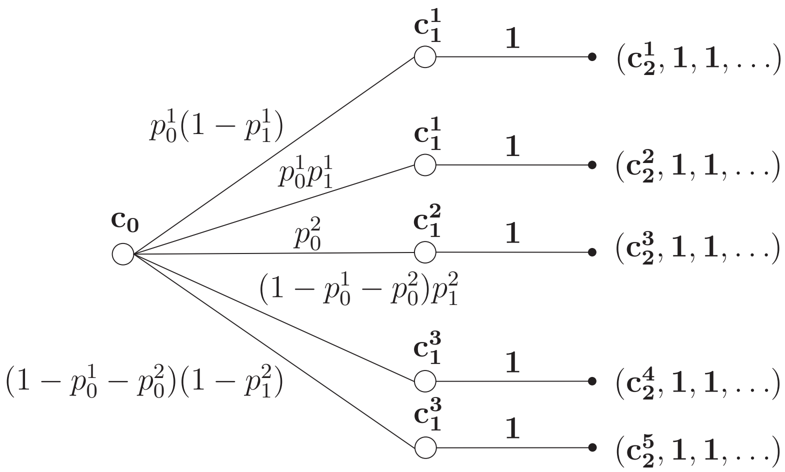

Figure 5: Trivial D_2 tree (example)

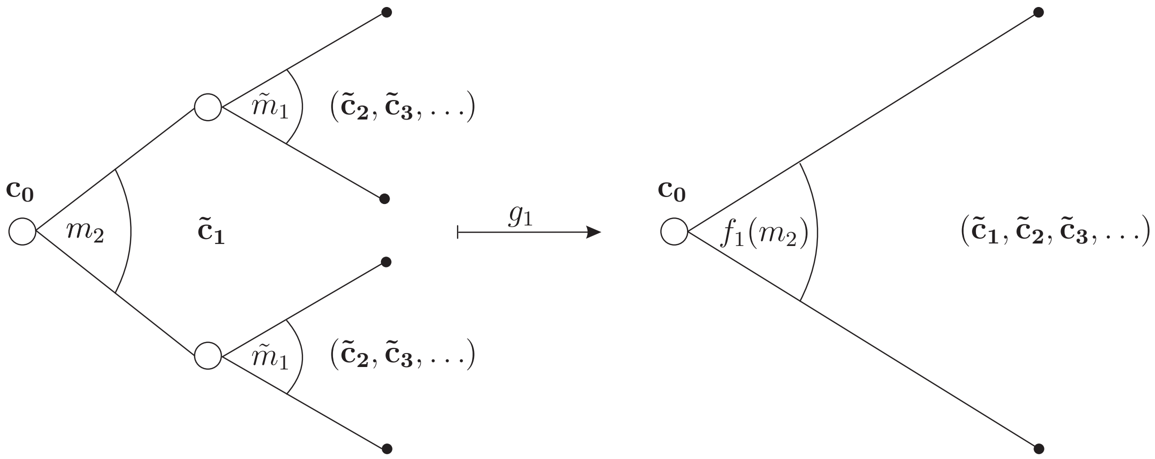

Figure 6: Consistent reduction

In general, consistent reduction can formally be stated as follows. Starting with the reduction of a two–stage probability tree to a tree of length 1, we first note that such a tree is given by $d_2 = (c_0, m_2)$, where $m_2 \in \mathcal{M}\left(\mathbb{R}_+ \times \mathcal{M}(\mathbb{R}_+^\infty)\right)$ is the probability measure over the random tuple $(\tilde{c}_1, \tilde{m}_1)$, i.e. over tomorrow's consumption level $\tilde{c}_1$ and the joint probability measure for $(\tilde{c}_2, \tilde{c}_3, \ldots)$. Hence, as exemplified above, for every $B \in \mathcal{B}(\mathbb{R}_+^\infty)$ we get

$$\begin{aligned} P(B) \equiv & P\left(\{(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \ldots) \in B\}\right) = \\ \int\limits_{\mathbb{R}_+ \times \mathcal{M}(R_+^\infty)} & \tilde{m}_1\left(\left\{(\tilde{c}_2, \tilde{c}_3, \ldots) \in \mathbb{R}_+^\infty | (\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \ldots) \in B\right\}\right) dm_2(\tilde{c}_1, \tilde{m}_1). \end{aligned}$$

Accordingly, we define

$$\begin{array}{llll} f_1: & \mathcal{M}(D_1) & \rightarrow & \mathcal{M}(D_0) \equiv \mathcal{M}(\mathbb{R}_+^\infty) \\ & m_2 & \mapsto & f_1(m_2): \quad \begin{array}{ll} \mathcal{B}(\mathbb{R}_+^\infty) & \rightarrow [0,1] \\ B & \mapsto \int \tilde{m}_1\left(\left\{y \in \mathbb{R}_+^\infty | (\tilde{c}_1, y) \in B\right\}\right) dm_2. \end{array} \end{array}$$

The mapping f_1 directly yields a probability measure over consumption sequences $(\tilde{c}_1, \tilde{c}_2, \ldots)$ out of a probability measure m_2 over nodes of one–stage trees. Thus, we define

$$\begin{array}{llll} g_1: & D_2 & \rightarrow & D_1 \\ & (c_0, m_2) & \mapsto & (c_0, f_1(m_2)). \end{array}$$

This is the desired mapping for consistently reducing a two–stage tree to its one–stage counterpart (cf. figure 6). Note again that, as illustrated in the example above, g_1 is not injective. One can furthermore show that f_1 and g_1 are both continuous and therefore measurable.

We now pursue inductively. Suppose the desired continuous mappings

$$\begin{array}{rccc} f_i: & \mathcal{M}(D_i) & \rightarrow & \mathcal{M}(D_{i-1}) \\ g_i: & D_{i+1} & \rightarrow & D_i \end{array}$$

have already been constructed for $i = 1, \ldots, t-1, t > 1$, such that $g_i(d_{i+1})$ is the resulting tree of length i that consistently corresponds to d_{i+1}. Consider an arbitrary probability tree $d_{t+1} \in D_{t+1}$ of length $t+1$. By definition, d_{t+1} is a tuple (c_0, m_{t+1}) of a non-negative real number c_0 representing today's consumption and a probability measure m_{t+1} over nodes of probability trees of length t,

$$d_{t+1} = (c_0, m_{t+1}) \in \mathbb{R}_+ \times \mathcal{M}(D_t).$$

Let $\tilde{d}_t$ denote the random variable of trees of length t emerging at stage 1 distributed according to m_{t+1}. If we want to consistently shorten the $(t+1)$–stage tree d_{t+1} by one step, we have to shorten the random t–stage tree $\tilde{d}_t$ by one step. Now, we already know that the latter reduction is executed via the mapping g_{t-1}. So by $g_{t-1}(\tilde{d}_t)$ we get the random $(t-1)$–stage tree emanating from the first node of the desired t–stage tree d_t that consistently corresponds to d_{t+1}. The wanted induced probability measure over $(t-1)$–stage trees is thus given by the distribution of $g_{t-1}(\tilde{d}_t)$, i.e. by

$$m_{t+1} \circ g_{t-1}^{-1} : \mathcal{B}_{t-1} \rightarrow [0,1].$$

We accordingly set

$$\begin{array}{rccc} f_t: & \mathcal{M}(D_t) & \rightarrow & \mathcal{M}(D_{t-1}) \\ & m_{t+1} & \mapsto & m_{t+1} \circ g_{t-1}^{-1} \end{array}$$

and

$$\begin{array}{rccc} g_t: & D_{t+1} & \rightarrow & D_t \\ & (c_0, m_{t+1}) & \mapsto & (c_0, f_t(m_{t+1}))\,. \end{array}$$

It follows inductively, that all f_t and g_t are continuous and thus measurable and also that g_t is not injective for all t. Observe that it is this non-injectivity that gives rise to the notion of nonindifference towards the timing of uncertainty resolution.

Now that we formally described what it means for a sequence of trees $(d_1, d_2, \ldots), d_t \in D_t$, to be consistent, we round off this subsection replicating a result in Epstein and Zin (1989) proving what we have already stated intuitively at the beginning of this section. Namely, every infinite probability tree can be identified unambiguously with a tuple of current consumption and a probability measure over nodes of probability trees emerging at period 1.

Definition 2.3.1. *We define $(D, \mathcal{B})$ as the inverse limit of the separable measurable spaces $(D_t, \mathcal{B}_t), t = 1, 2, \ldots$, relative to the measurable mappings $g_t : D_{t+1} \to D_t$. I.e.*

(i) $D = \{(d_1, d_2, \ldots) \in \prod_{t=1}^{\infty} D_t | d_t = g_t(d_{t+1})$ *f.a.* $t \geq 1\}$

(ii) *$\mathcal{B}$ is the smallest σ–algebra of subspaces of D that renders the canonical projection*

$$\pi_t : D \to D_t, (d_1, d_2, \ldots) \mapsto d_t$$

measurable for all $t = 1, 2, \ldots$, i.e. $\mathcal{B} = \sigma\left(\bigcup_{t=1}^{\infty} \{\pi_t^{-1}(B_t) | B_t \in \mathcal{B}_t\}\right)$.

Theorem 2.3.1. *First, D is a Polish space relative to the subspace topology that is induced by the product topology on $\prod_{t=1}^{\infty} D_t$ and $\mathcal{B}$ equals the Borel σ–algebra that is generated by this topology on D. Second, D is homeomorphic to $\mathbb{R}_+ \times \mathcal{M}(D)$.*

Proof. The first part follows from Parthasarathy (1967), Theorem 2.6. For the second part, let $d = (d_1, d_2, \ldots) \in D, d_t = (c_0, m_t) \in D_t, d_t = g_t(d_{t+1})$ f.a. $t \in \mathbb{N}$ be arbitrary. Thus, by definition we have for all $t \in \mathbb{N}$ a probability measure m_{t+1} on $\mathcal{B}_t$ satisfying $m_{t+1} = f_{t+1}(m_{t+2}) = m_{t+2} \circ g_t^{-1}$. Following Parthasarathy (1967), Theorem 3.2, there exists a unique probability measure $m : \mathcal{B} \to [0, 1]$ such that $m\left(\pi_t^{-1}(B_t)\right) = m_{t+1}(B_t)$ f.a. $t \in \mathbb{N}, B_t \in \mathcal{B}_t$. By setting $\Phi(d) := (c_0, m)$ we define the mapping $\Phi : D \to \mathbb{R}_+ \times \mathcal{M}(D)$. One can now show that this mapping is a homeomorphism. □

Before we proceed, let us provide some additional intuition about the constructed homeomorphism and its significance for the upcoming utility analysis. Consider again an arbitrary probability tree $d = (d_1, d_2, \ldots) \in D, d_t = (c_0, m_t) \in D_t, d_t = g_t(d_{t+1})$ f.a. $t \in \mathbb{N}$. Approximated until stage

$(t+1)$, this tree is given by $\pi_{t+1}(d) = d_{t+1} = (c_0, m_{t+1})$ and the trees of length t emerging at the first stage of d_{t+1} are distributed according to m_{t+1}. Yet on the other hand, the probability trees originating at the first stage of d_{t+1} also correspond to the t–stage approximations of the infinite probability trees emerging at the first stage of the whole tree d. Since the infinite probability trees emerging at stage 1 of d are distributed according to the probability measure m and their t–stage approximation is given via the mapping π_t, these t–stage approximations are distributed according to the probability distribution $m \circ \pi_t^{-1}$. So it must hold that

$$m \circ \pi_t^{-1} = m_{t+1}.$$

Moreover, observe that the fact that D is homeomorphic to $\mathbb{R}_+ \times \mathcal{M}(D)$ importantly says that for every tree $d \in D$ it holds that every "subtree" that emanates from some of its intermediary nodes necessarily also lies in D. This "stationarity" of the consumption space is necessary for the existence of a recursive utility function on D.[8]

2.3.3 Additional restrictions to the lottery space

The space D will provide the fundament for our consumption space, i.e. the space over which decisions will be made. In order to describe such decisions, we will introduce eligible utility functions in the next section. However, these utility functions can generally not be defined on the whole space D but only on particular subspaces. In fact, consumption has to be bounded in some sense. Therefore, we have to further narrow the lottery space appropriately.

For that purpose, Epstein and Zin (1989) define the space of consumption sequences such that the gross growth rate is capped by some $b \geq 1$, i.e.

$$Y(b,l) := \left\{ (c_0, c_2, \ldots) \in \mathbb{R}_+^\infty \middle| \frac{c_t}{b^t} \leq l \text{ f.a. } t \in \mathbb{N} \right\} = \prod_{t=0}^{\infty} [0, b^t l].$$

Endow $Y(b,l)$ with the product topology and note that according to Tychonoff's theorem $Y(b,l)$ is compact. The subspaces $D(b)$ of D consisting only of probability trees $d = (d_1, d_2, \ldots)$, $d_t = (c_0, m_t)$, such that the atemporal probability measure m_1 gives rise to consumption sequences $(c_1, c_2, \ldots)$ in $Y(b,l)$ for some $l > 0$ with probability 1, are suitable as domains for recursive utility functions.

[8]Cf. Epstein and Zin (1989), p. 941, and the introductory remarks to section 3.

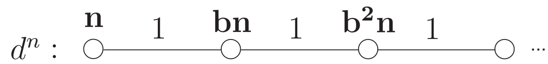

Figure 7: Degenerate temporal lottery

Definition 2.3.2. *For $b \geq 1$ define*

$$D(b) := \{d = (d_1, d_2, \dots) \in D | d_1 = (c_0, m_1) \text{ s.t. } \exists l > 0 : \text{supp}(m_1) \subset Y(b, l)\}$$

and endow $D(b)$ with the subspace topology.

Note that since $D(b)$ is a subspace of a separable metric space, it is thus a separable metric space itself. Moreover, as a subset of D it is homeomorphic to a subset of $\mathbb{R}_+^\infty \times \mathcal{M}(D)$, i.e. via Φ every probability tree $d \in D(b)$ can uniquely be represented as a tuple of the consumption level c_0 today and a probability measure over nodes of trees emanating at period 1. However, as the next example will illustrate, not every probability measure $m \in \mathcal{M}(D(b))$ is in question for this identification.

Example 3 For $n \geq 1$ consider the probability tree d^n depicted in figure 7. The corresponding 1–stage trees are given by $d_1^n = (n, m_1^n) \in D_1$, where $m_1^n = \delta_{(bn, b^2n, b^3n, \dots)}, n \geq 1$. Since $\text{supp}(m_1^n) \subset Y(b, n), n \geq 1$, it holds that $d^n \in D(b)$ f.a. $n \geq 1$. But now consider the probability defined by

$$m = \sum_{n=1}^{\infty} \frac{1}{2^n} \delta_{d^{bn}} \in \mathcal{M}(D(b)).$$

We want to show that the tree $d \in D$ that via Φ corresponds to the tuple $(1, m) \in \mathbb{R}_+ \times \mathcal{M}(D(b))$ does not lie in $D(b)$. The tree is pictured in figure 8 and it can be seen that its 1-stage approximation is given by $d_1 = (1, m_1)$, where $m_1 = \sum_{n=1}^{\infty} \frac{1}{2^n} \delta_{(bn, b^2n, b^2n, \dots)}$. Since for every $b \geq 1$ there is no $l > 0$ such that $\text{supp}(m_1) \subset Y(b, l)$, it follows that $d \notin D(b)$.

This example makes clear that $D(b)$ cannot be homeomorphic to $\mathbb{R}_+ \times \mathcal{M}(D(b))$ but that only a subset of $\mathcal{M}(D(b))$ is in question. In order to see which probability measures these are, let us again consider an arbitrary tuple $(c_0, m) \in \mathbb{R}_+ \times \mathcal{M}(D(b))$. The corresponding tree $d = (d_1, d_2, \dots) \in D, d_t = (c_0, m_t)$, is identified through $m_{t+1} = m \circ \pi_t^{-1}, t \geq 1$ (cf. Theorem 2.3.1). In particular, we have that $m_2 = m \circ \pi_1^{-1}$ and hence $m_1 = f_1(m \circ \pi_1^{-1})$. In order

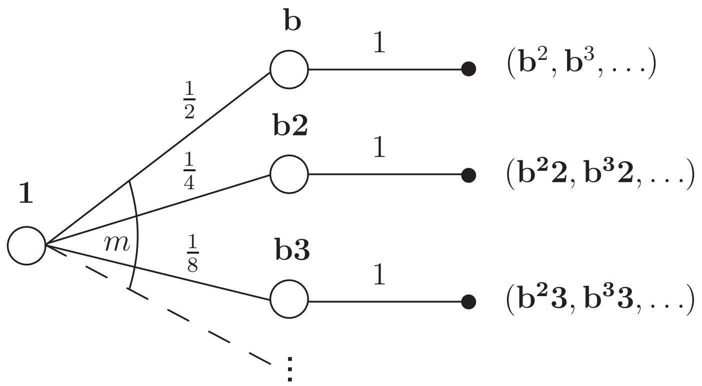

Figure 8: Temporal lottery *not* in $D(b)$

for d to lie in $D(b)$ it is necessary and sufficient that $\operatorname{supp}(f_1(m \circ \pi_1^{-1})) \subset Y(b,l)$ for some $l > 0$. If we restrict the homeomorphism $\Phi : D \to \mathbb{R}_+ \times \mathcal{M}(D)$ to $D(b)$ we get a homeomorphism

$$\Phi\big|_{D(b)} : D(b) \to \mathbb{R}_+ \times \{m \in \mathcal{M}(D(b)) | \operatorname{supp}(f_1(m \circ \pi_1^{-1})) \subset Y(b,l), \text{ for some } l > 0\}.$$

We finish this section by summarizing these considerations with the following Theorem.

Theorem 2.3.2. *$D(b)$ is homeomorphic to $\mathbb{R}_+ \times \hat{\mathcal{M}}(D(b))$ with*

$$\hat{\mathcal{M}}(D(b)) := \{m \in \mathcal{M}(D(b)) | \operatorname{supp}(f_1(m \circ \pi_1^{-1})) \subset Y(b,l), \text{ for some } l > 0\},$$

where $\hat{\mathcal{M}}(D(b))$ is endowed with the subspace topology induced from $\mathcal{M}(D(b))$.

3 EZ utility

Having reached a formal description of the concept of an infinite probability tree, we next want to describe the decision making over such trees by means of the EZ utility representation. Thereby, we parallel the approach taken above in that we identify the decision over such infinite trees with the two-stage decision over current consumption and future utility prospects. I.e., precisely how we represented an infinite probability tree for the infinite random consumption sequence $(c_0, c_1, c_2 \ldots)$ by a pair (c_0, m) of period 0 consumption and a probability measure over nodes of probability trees emanating next period, we now identify the utility of such an infinite consumption sequence $(c_0, c_1, c_2 \ldots)$ by a recursive two–period utility stemming from today's consumption level c_0 and future utility contingent on next period's random node.

The mechanics of an EZ utility function $U : D(b) \to \mathbb{R}_+$ can be described as follows. Let $d \in D(b)$ denote an arbitrary infinite probability tree identified with the tuple $(c_0, m) \in \mathbb{R}_+ \times \hat{\mathcal{M}}(D(b))$ via the above constructed homeomorphism (cf. Theorem 2.3.2). Hence, the random node of a probability tree $\tilde{d}$ emerging at period 1, is distributed according to the probability measure m. Thus, if we evaluate possible next period nodes by their utility through $U(\tilde{d})$, we get an induced probability distribution over random real utility levels as of tomorrow. Given measurability of U, this random future utility is distributed according to

$$m_U := m \circ U^{-1} : \mathcal{B}(\mathbb{R}_+) \to [0, 1].$$

In other words, the probability measure m_U describes an atemporal lottery over next period's random utility. Given monotonic preferences over such atemporal utility lotteries,[9] we will assume these to be representable through a certainty equivalence functional

$$\mu : \mathcal{M}(\mathbb{R}_+) \to \mathbb{R}_+,$$

meaning for any lottery $P \in \mathcal{M}(\mathbb{R}_+)$ the decision maker is indifferent between the lottery itself and a certain utility level $\mu(P)$. Note that μ thereby aggregates the decision maker's evaluation of uncertain future utility along next period's nodes. This perspective gives rise to the intuitive notion of uncertainty aggregation as coined by Träger (2011). Finally, this certainty equivalent is combined with today's consumption c_0 via

$$W : \mathbb{R}_+^2 \to \mathbb{R}_+,$$

acting as a time aggregator of the utility contributions of the elements of (c_0, m). To sum up, for all $d \in D(b)$

$$U(d) = W\left(\Phi_1(d), \mu\left(\Phi_2(d) \circ U^{-1}\right)\right) = W(c_0, \mu(m_U)). \tag{1}$$

[9]By assuming monotonic preferences we plausibly model the decision maker as favoring higher utility levels over lower, but see later.

A utility function U over $D(b)$ that satisfies the above equation is called recursive. Koopmans (1960) introduced the notion of a recursive utility function on infinite deterministic consumption programs by means of a function that aggregates current consumption and prospects of future continuation utility. On the other hand, Selden (1978) provided a representation for two-period "certain$\times$uncertain" consumption programs, where utility is defined as an aggregation of current consumption and a certainty equivalent of next period's random consumption level. A time-consistent multi-period extension of the Selden representation is found in Kreps and Porteus (1978).[10] Finally, the stochastic complement to the infinite horizon recursive Koopmans representation is eventually provided by Epstein and Zin (1989) as described above. Note thereby that assuming an EZ representation is sufficient for the underlying preference ordering over temporal consumption lotteries to satisfy stationarity and intertemporal consistency.[11] Further note the relation of such stationarity of preferences and the "stationarity" of the consumption space as proven above.

3.1 Uncertainty aggregation

As already sketched above, the idea of a certainty equivalent with regard to random continuation utility is to assign a deterministic level of appreciation (utility level) to every atemporal lottery over continuation utilities, which renders the decision maker indifferent to the latter. In order for this to make sense we first have to introduce preferences over atemporal continuation utility lotteries. Thereby, we discuss a set of assumptions about such preferences and their respectively implied utility representations. The employed axiomatization is provided by Chew (1989). It is flexible enough to account for prominent behavioral peculiarities that have been reported as violations of the classic von Neumann/Morgenstern (vNM) independence axiom.[12]

Thereafter, we will specify the functional form of certainty equivalents that come with such utility representations. We conclude with two parametric examples of particular interest.

[10] Cf. the introductory remarks in Weil (1990).

[11] Cf. Epstein and Zin (1989), p. 945.

[12] See e.g. the Allais (1953) paradox as the most prominent of such violations.

3.1.1 Preferences over utility lotteries

Let $I \subset \mathbb{R}_+$ denote a compact interval and again write $\mathcal{B}(I)$ and $\mathcal{M}(I)$. We assume that preferences over lotteries in $\mathcal{M}(I)$ are given by a binary relation $\preceq$. Further let $\prec$ and $\sim$ denote the induced strict preference relation and indifference relation, respectively. Consider the following behavioral axioms:

(O) $\preceq$ is a weak order, i.e. complete and transitive.

(C) $\forall P \in \mathcal{M}(I) : \{Q \in \mathcal{M}(I) | P \prec Q\}$ and $\{Q \in \mathcal{M}(I) | Q \prec P\}$ are open with respect to the weak topology on $\mathcal{M}(I)$.

(VWS) $\forall P, Q \in \mathcal{M}(I)$:

$$P \sim Q \Rightarrow \forall R \in \mathcal{M}(I), \lambda \in (0,1) : \\ \exists \theta \in (0,1) : \lambda P + (1-\lambda)R \sim \theta Q + (1-\theta)R.$$

(WS) $\forall P, Q \in \mathcal{M}(I)$:

$$P \sim Q \Rightarrow \forall \lambda \in (0,1) : \exists \theta \in (0,1) : \\ \forall R \in \mathcal{M}(I) : \lambda P + (1-\lambda)R \sim \theta Q + (1-\theta)R.$$

(S) $\forall P, Q, R \in \mathcal{M}(I)$:
$P \sim Q \Rightarrow \forall \lambda \in (0,1) : \lambda P + (1-\lambda)R \sim \lambda Q + (1-\lambda)R.$

Weak order (O) and continuity (C) are the standard requirements for there to exist a continuous numerical function representing $\preceq$ over $\mathcal{M}(I)$.[13] We will thus always demand preferences to satisfy (O) and (C). Assuming $\preceq$ to also obey one of the remaining three axioms imposes considerably more structure on the decision making. Thereby, (VWS), (WS) and (S) increasingly facilitate the analytical implementation of the resulting utility representation in applied work. Nevertheless, their empirical appeal is decreasing in the same order along their respective degree of reconcilability with behavioral data.[14]

Starting with the most restrictive axiom, substitution (S) demands that whenever the decision maker is indifferent between two lotteries P and Q, he

[13] Cf. Debreu (1954).

[14] See how Chew (1983) motivates his inquiry into utility representations over lottery spaces that do not necessarily satisfy the classic independence axiom, which is rephrased to (S) in the Chew (1989) framework.

is also indifferent between the mixture of P with a third lottery R and the mixture of Q with that third lottery R, both mixtures by the same ratio.

Now, weak substitution (WS) weakens (S) as it additionally permits the mixture ratios at which the decision maker is indifferent between a mixture of P or Q with a third lottery R to differ. Yet, these mixtures cannot depend on the third lottery R. In other words, it is possible that one of the two indifferent lotteries can be mixed more "easily" with the third leaving room for complementarity between P and R or Q and R.

Eventually, with very weak substitution (VWS), the ratio at which indifference is attained is additionally allowed to depend on R. Hence, (S) implies (WS), which in turn implies (VWS).

3.1.2 Chew certainty equivalents

General representation Chew (1989), Theorem 3, proves necessity and sufficiency of the "very weak" constellation of the above axioms for there to exist a continuous utility representation of $\preceq$ over $\mathcal{M}(I)$ of the so called implicit-weighted kind. Thereby, the utility level $u(P)$ associated with a lottery $P \in \mathcal{M}(I)$ is the unique root of a mapping

$$y \mapsto \int_I \Psi(x,y)dP(x),$$

where $\Psi : I \times \mathbb{R} \to \mathbb{R}$ has to satisfy certain continuity requirements. Implicitly defining the continuous weight function $w : I \times \mathbb{R} \to \mathbb{R}$ such that

$$\Psi(x,y) = w(x,y)(v(x) - y),$$

where the continuous mapping $v : I \to \mathbb{R}$ is defined via $v(x) := u(\delta_x)$,[15] thus yields

$$0 = \int_I \Psi(x,u(P))dP(x) = \int_I w(x,u(P))v(x)dP(x) - u(P)\int_I w(x,u(P))dP(x),$$

such that

$$u(P) = \int_I \frac{w(x,u(P))}{\int_I w(x,u(P))dP(x)} v(x)dP(x). \tag{2}$$

In a sense, the utility u associated with a lottery $P \in \mathcal{M}(I)$ is given by a weighted expected value of the utility index v over deterministic outcomes.

[15] I.e. $v(\cdot)$ is a utility index over deterministic outcomes in I.

The fact that the weights in this representation also depend on the utility level $u(\cdot)$ of the considered lottery itself gives rise to the notion of implicit weighting.[16]

Next, since our analysis is concerned with lotteries over utility levels, the preference ordering $\preceq$ is also assumed to always satisfy the following strict monotonicity axiom:

(M) For all $x, y \in I$ it holds that $\delta_x \prec \delta_y \Leftrightarrow x < y$.

Hence, v is strictly monotonically increasing on I and therefore invertible so that for every $P \in \mathcal{M}(I)$ there is a unique real number in I, denoted by $\mu(P)$ and called lottery P's certainty equivalent, that satisfies

$$P \sim \delta_{\mu(P)} \Leftrightarrow u(P) = u(\delta_{\mu(P)}) = v(\mu(P)). \tag{3}$$

Hence, by (M), preferences over $\mathcal{M}(I)$ can equivalently be stated in terms of certainty equivalents, i.e. for all $P, Q \in \mathcal{M}(I)$ it holds that

$$P \preceq Q \Leftrightarrow \mu(P) \leq \mu(Q).$$

With regard to the notion of uncertainty aggregation, note that μ makes explicit its ingredients. Next to the description of uncertainty $P \in \mathcal{M}(I)$, μ incorporates both the appreciation of deterministic continuation utility levels $v(\cdot)$ and the respective implicit weights of the utility representation.

Next, (3) implies that

$$\begin{aligned} &\int_I \Psi(x, v(\mu(P)))dP(x) = 0 \\ \Leftrightarrow &\int_I \psi(x, \mu(P))dP(x) = 0, \end{aligned}$$

where $\psi : I \times I \to \mathbb{R}, (x, y) \mapsto \Psi(x, v(y))$. I.e., given preferences $\preceq$ over $\mathcal{M}(I)$ which satisfy (O), (C), (VWS) and (M), the certainty equivalent $\mu(P)$ of a lottery $P \in \mathcal{M}(I)$ is given by the unique root of the mapping

$$y \mapsto \int_I \psi(x, y)dP(x).$$

[16] It is the particular functional form of w that comprises behavioral implications about the underlying decision making, et vice versa, but see shortly.

Note that by construction we have

$$\psi(x,x) \equiv \Psi(x,v(x)) = \int_I \Psi(\tilde{x},v(x))d\delta_x(\tilde{x}) = \int_I \Psi(\tilde{x},u(\delta_x))d\delta_x(\tilde{x}) = 0,$$

Thus, to put it explicitly, the considered certainty equivalents satisfy the "consistency with certainty" property listed by Chew (1983) as a crucial requirement for mean value functionals, i.e.

$$\mu(\delta_x) = x.$$

Moreover, Chew (1989), Theorem 5, proves that such certainty equivalents are consistent with first (resp. second) degree stochastic dominance if and only if for all $y \in I$ the mapping $x \mapsto \psi(x,y)$ is strictly monotonically increasing (resp. concave).

Eventually, if we assume preferences to have a Chew-type utility representation over $\mathcal{M}(I)$, with $I = \mathbb{R}_+$, we further restrict the certainty equivalents to satisfy homogeneity:

(H) For all $P \in \mathcal{M}(I)$ and $\lambda > 0$ it holds that $\mu(P_\lambda) = \lambda\mu(P)$,

where P_λ denotes the probability measure defined by $P_\lambda(B) := P(x \in I | \lambda x \in B)$. Intuitively, (H) requires the assignment of a λ-fold certainty equivalent to a lottery that yields a λ-fold utility. This implies

$$0 = \int_I \psi(x,\mu(P_\lambda))dP_\lambda(x) = \int_I \psi(\lambda x,\lambda\mu(P))dP(x),$$

which is satisfied in particular for ψ linear homogenous. Thus, by defining $\zeta(x) := \psi(x,1)$, we find the sought for Chew (1989)-type certainty equivalent $\mu(P)$ of a continuation utility lottery $P \in \mathcal{M}(I)$ to be the unique root of the mapping

$$y \mapsto \int_I \zeta\left(\frac{x}{y}\right) dP(x).$$

Special cases We finish this subsection on Chew-type implicit-weighted certainty equivalents for some utility lottery $P \in \mathcal{M}(I)$ by presenting two parametric examples explicitly considered by Epstein and Zin (1989).

First, let $\zeta(x) = \frac{x^\alpha - 1}{\alpha} + a(x-1), 0 \neq \alpha < 1, a \geq 0$. This gives rise to a so called Chew/Dekel (CD) certainty equivalent implicitly defined as the solution to

$$\mu_{CD}(P)(1+a\alpha) - a\alpha \int_I x dP(x)) = (\mu_{CD}(P))^{1-\alpha} \int_I x^\alpha dP(x). \tag{4}$$

Second, consider the case of $a = 0$ in the above example, which allows us to explicitly solve for the certainty equivalent,

$$\mu_{KP}(P) = \left(\int_I x^\alpha dP(x) \right)^{\frac{1}{\alpha}}. \tag{5}$$

This is the certainty equivalent of a decision maker with so called Kreps/Porteus (KP) preferences. Observe by (2) that this μ-specification arises in the case of a constant (explicit) weight function w and for $v(x) = x^\alpha$, $0 \neq \alpha < 1$. To put it another way, next to a constant relative risk aversion (CRRA)-type utility index $v(\cdot)$, such a representation demands the decision maker to obey (O), (C) and (S) by the vNM Theorem.[17]

3.2 Time aggregation

So far, we saw how the notion of recursive utility "reduces" the problem of evaluating intertemporal consumption tradeoffs within an infinite horizon framework to a two–period problem of assessing the tradeoff between current consumption and a certainty equivalent of random utility prospects. In order to merge these two ingredients we already introduced the time aggregator

$$W : \mathbb{R}^2_+ \to \mathbb{R}_+.$$

Epstein and Zin (1989) explicitly demand this aggregator to have the form

$$W(c, \mu) = [c^\rho + \beta\mu^\rho]^{\frac{1}{\rho}}, \; 0 \neq \rho < 1, \beta \in (0, 1). \tag{6}$$

Note that it is made sure that both today's consumption and future utility enter the modelled decision maker's evaluation positively.

To complement our analysis of parametric examples of Chew certainty equivalents above, finally consider the special case of the KP functional further restricted to $\alpha = \rho$, where ρ is the time aggregation parameter above

$$\mu_{EU}(P) := \left(\int_I x^\rho dP(x) \right)^{\frac{1}{\rho}}. \tag{7}$$

[17]See Chew (1989) Theorem 1 for a formal statement of the vNM Theorem in this context.

Note that Chew (1989), Theorem 2, also proves necessity and sufficiency of the "weak" constellation for a "(explicit) weighted utility" representation of intermediate generality. We skip this part because we will not refer to it in our ensuing analysis. The (WS) axiom is thus listed merely for the sake of exposition.

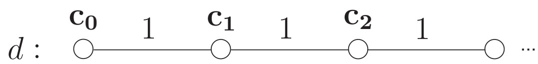

Figure 9: Deterministic consumption sequence

This specification finally gives rise to an expected utility (EU) representation over temporal lotteries and thus yields the standard model as described in the introductory remarks to section 2.

Let us next restrict our attention to probability trees that correspond to deterministic consumption sequences $(c_0, c_1, c_2, \ldots)$, such as the tree d depicted in figure 9. By

$$\begin{aligned} d &\cong (c_0, m), \quad \text{where } m = \delta_{d^1} \\ d^1 &\cong (c_1, m^1), \quad \text{where } m^1 = \delta_{d^2} \\ d^2 &\cong (c_2, m^2), \quad \text{where } m^2 = \delta_{d^3} \\ &\text{etc.} \end{aligned}$$

it follows from (1) that

$$\begin{aligned} U(d) &= [c_0^\rho + \beta\mu(\delta_{U(d^1)})^\rho]^{\frac{1}{\rho}} = \\ &= [c_0^\rho + \beta U(d^1)^\rho]^{\frac{1}{\rho}} = \\ &= [c_0^\rho + \beta[c_1^\rho + \beta\mu(\delta_{U(d^2)})^\rho]^{\frac{1}{\rho}\rho}]^{\frac{1}{\rho}} = \\ &= [c_0^\rho + \beta c_1^\rho + \beta^2 U(d^2)^\rho]^{\frac{1}{\rho}} = \ldots = \\ &= \left[\sum_{t=0}^{\infty} \beta^t c_t^\rho\right]^{\frac{1}{\rho}}. \end{aligned} \tag{8}$$

I.e., the approach taken here results in a utility function of the constant elasticity of substitution (CES) class as an evaluator of *deterministic* consumption sequences.[18] In this context, the parameter

$$\text{EIS} := \frac{1}{1-\rho}$$

is referred to as the elasticity of intertemporal substitution of consumption.

[18]Observe that, since $\text{id}^{\frac{1}{\rho}}$ is a strictly monotonically increasing transformation, a standard Samuelson (1937) Discounted Utility function $\sum_{t=0}^{\infty} \beta^t c_t^\rho$ is an alternative representation of the underlying preferences if restricted to degenerate temporal lotteries.

We conclude with the following Theorem that ensures existence of the considered EZ utility functions.[19]

Theorem 3.2.1. *If W has the CES form (6), then, for the three parametric examples of μ considered above, the functional equation*

$$U(c_0, m) = W(c_o, \mu(m_U))$$

has a solution

1. *for $\rho > 0 : V : D(b) \to \mathbb{R}_+,$ where b satisfies $\beta b^\rho < 1$*
2. *for $\rho < 0 : V : D \to \mathbb{R}_+$.*

3.3 Timing and risk preferences

The major advantage of adopting the more general EZ utility for applied work stems from the disentanglement (however incomplete) of the decision maker's attitude towards risk and towards the timing of consumption. This subsection demonstrates their separation.

3.3.1 Definitions

We begin by defining the notion of timing and risk preferences, respectively. Both definitions are mutually abstract in the following sense. In defining timing preferences we abstract from uncertainty, while in defining risk preferences we keep the analysis atemporal.

Moreover, we are particularly interested in a comparative assessment of preferences. We therefore consider two recursive decision makers with utility representations U^I and U^{II}. Specifically, for $i = I, II$, $0 \neq \rho^i < 1$ and $0 < \beta^i < 1$

$$V^i(c_0, m) = W^i\left(c_0, \mu^i(V_m^i)\right) = \left[c_0^{\rho^i} + \beta^i\left(\mu^i(V_m^i)\right)^{\rho^i}\right]^{\frac{1}{\rho^i}},$$

where $\mu^i : \mathcal{M}(\mathbb{R}_+) \to \mathbb{R}_+$ is one of the three parametric examples of the Chew class studied above.

[19] Epstein and Zin (1989) prove a more general version in their Theorem 3.1.

Timing preferences The deterministic treatment of consumption timing and the associated intertemporal consumption substitution is at least two-dimensional in (modern) economics.[20] First, there is mere impatience such that a decision maker c.p. favors current consumption. Second, he also assesses the "relative abundance" of consumption available to him over time.[21]

Now, impatience is formally reflected in the discount parameter β. We thus regard the decision maker V^I as less patient than V^{II} if he discounts future consumption more strongly, i.e. if it holds that

$$\beta^I \leq \beta^{II}.^{22} \tag{9}$$

As far as deterministic consumption substitution is concerned, we note that, by means of the utility U, the EIS controls the decision maker's affinity toward a smooth consumption profile. We thus interpret V^I to be more averse toward deviations from smooth consumption than V^{II} if

$$\text{EIS}^I := \frac{1}{1-\rho^I} \leq \frac{1}{1-\rho^{II}} =: \text{EIS}^{II},$$

or equivalently

$$\rho^I \leq \rho^{II}. \tag{10}$$

Risk preferences As opposed to timing preferences, a readily operationalized measure of comparative risk aversion is less evident. However, it seems natural to focus on the employed certainty equivalent functional as it serves as the uncertainty aggregation device in the EZ framework. Accordingly, for $W^I = W^{II}$, Epstein and Zin (1989) define V^I to be more risk averse than V^{II} if it holds that

$$\mu^I \leq \mu^{II}.^{23} \tag{11}$$

[20]A thorough discussion on the notion of time preference, its historic development and also some serious reservations against discounted utility models is provided by Frederick, Loewenstein, and O'Donoghue (2002).

[21]Cf. Fisher (1930), p. 67.

[22]Note that the notion of comparative impatience can be defined more formally in terms of preferences along the lines of Olson and Bailey (1981) by means of the decision makers' marginal rate of substitution between consumption levels in two different periods after excluding "the effect of a difference in marginal utility."

[23]Note that, as remarked by Epstein and Zin (1989), for the case of KP certainty equivalents, by the assumed consistency with second order stochastic dominance it follows that the least risk averse decision maker aggregates uncertainty additionally obeying $\alpha = 1$. I.e. his utility index satisfies $v(\cdot) = \text{id}(\cdot)$ such that his certainty equivalent is a plain expected value.

3.3.2 Disentangling attitudes towards risk and timing

Separation By further examining the above definitions we first find for the CD-class that (11) is equivalent to

$$\alpha^I \leq \alpha^{II} \text{ and } a^I \leq a^{II}.$$

Second, for the KP-class ($a^I = a^{II} = 0$) this condition reduces to

$$\alpha^I \leq \alpha^{II}.$$

Thus, the EZ utility representation with CD- and KP-certainty equivalents yields a parametric disentanglement of comparative risk aversion and timing preferences.[24]

Eventually, with the additional restriction of $\alpha^i = \rho^i, i = I, II$, it is obvious that such separation is impossible in the case of EU certainty equivalents.

Nonindifference towards the timing of uncertainty resolution Even though the EZ representation allows for a parametric disentanglement of the above notions of risk and timing preferences, it is important to note that this separation is only partly in nature. Specifically, departing from an EU certainty equivalent necessarily gives rise to nonindifference towards the timing of the resolution of consumption uncertainty as it is illustrated in the introductory example to the present paper.[25]

Precisely, Epstein and Zin (1989) conclude in the sense of Kreps and Porteus (1978), Theorem 3, that a decision maker V^i with KP preferences over atemporal continuation utility lotteries prefers earlier (later) resolution if and only if it holds that $\alpha^i < (>) \rho^i$. Moreover, V^i is indifferent to the timing of uncertainty resolution if and only if it holds that $\alpha^i = \rho^i$, i.e. if and only if he has EU preferences.

[24]Note that this separation as well as the utility representation is akin to Selden (1978)'s result in a two period environment.

[25]Note that looked at it this way, such nonindifference may appear as a cost of the achieved disentanglement, cf. Epstein, Farhi, and Strzalecki (2014). Interestingly, as mentioned before, Kreps and Porteus (1978), who provided the fundament of the EZ framework and thus for the studied separation of risk aversion and EIS, were looking for a temporal utility representation that allowed for the explicit modelling of such nonindifference and did not motivate there analysis through the issue of entangled risk aversion and consumption substitutability. Nevertheless, they already indicated on some relation between risk aversion and nonindifference, see Kreps and Porteus (1978), p. 198.

3.4 The EZ/KP representation

We conclude this section on the EZ representation by the application of the KP case to our introductory example. The focus on KP preferences is natural in the sense that our ensuing analysis is focused on applied macroeconomics. For such work it is important to have an explicit functional form of U to parameterize.

Summing up, the EZ/KP utility representation reads

$$U(d) = \left[(\Phi_1(d))^\rho + \beta\left(\mathbb{E}_{\Phi_2(d)}\left[U^\alpha\right]\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{1-\rho}}, 0 \neq \rho < 1, 0 \neq \alpha < 1, \beta \in (0,1).^{26} \tag{12}$$

To illustrate its application, consider Robinson's two alternatives as presented in example 1 and compare his behavior for the three cases of him preferring earlier or later resolution or being indifferent about the timing of uncertainty resolution, i.e. following the consequentialist hypothesis implied by the standard model.

Example 1 (continued) Robinson's preferences over temporal lotteries are assumed to be representable by an EZ utility function of the KP form. He has to choose between the two (finite) temporal lotteries displayed in figure 1, i.e. between $d = (10, m)$ and $\hat{d} = (10, \hat{m})$, with $d, \hat{d} \in D$. Since Robinson decides in favor of the temporal lottery that results in a relatively higher utility evaluation, his decision making can be described as comparing

$$\begin{aligned} U(d) &= \left[10^\rho + \beta\mu(m \circ U^{-1})^\rho\right]^{\frac{1}{\rho}} \\ &= \left[10^\rho + \beta\left(\mathbb{E}_m\left[U^\alpha\right]\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}}, \beta \in (0,1), 0 \neq \rho < 1, \end{aligned} \tag{13}$$

and

$$\begin{aligned} U(\hat{d}) &= \left[10^\rho + \beta\mu(\hat{m} \circ U^{-1})^\rho\right]^{\frac{1}{\rho}} \\ &= \left[10^\rho + \beta\left(\mathbb{E}_{\hat{m}}\left[U^\alpha\right]\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}}, \beta \in (0,1), 0 \neq \rho < 1. \end{aligned} \tag{14}$$

Now, note again that m is already a lottery over degenerate trees (deterministic consumption sequences) while $\hat{m}$ is a degenerate lottery $\delta_{\hat{d}_1}$, with

[26] Note that Weil (1990)'s "generalized isoelastic" utility provides an equivalent representation.

$\hat{d}_1 = (10, \hat{m}_1)$. Moreover, $\hat{m}_1$ is a lottery over degenerate trees. We thus compute

$$\begin{aligned}
\mu(m \circ U^{-1}) &= \left(\frac{1}{2}U^{\alpha}(10, 10, \ldots) + \frac{1}{2}U^{\alpha}(10, 5, 5, \ldots)\right)^{\frac{1}{\alpha}} \\
&\overset{(8)}{=} \left(\frac{1}{2}\left(\left[\sum_{t=0}^{\infty} \beta^t 10^{\rho}\right]^{\frac{1}{\rho}}\right)^{\alpha} + \frac{1}{2}\left(\left[10^{\rho} + \sum_{t=1}^{\infty} \beta^t 5^{\rho}\right]^{\frac{1}{\rho}}\right)^{\alpha}\right)^{\frac{1}{\alpha}} \\
&= \left(\frac{1}{2}\left(10\left[\frac{1}{1-\beta}\right]^{\frac{1}{\rho}}\right)^{\alpha} + \frac{1}{2}\left(\left[10^{\rho} + 5^{\rho}\frac{\beta}{1-\beta}\right]^{\frac{1}{\rho}}\right)^{\alpha}\right)^{\frac{1}{\alpha}}
\end{aligned}$$

and

$$\begin{aligned}
\mu(\hat{m} \circ U^{-1}) &= U(\hat{d}_1) \\
&= \left[10^{\rho} + \beta\mu(\hat{m}_1 \circ U^{-1})^{\rho}\right]^{\frac{1}{\rho}} \\
&= \left[10^{\rho} + \beta\left(\frac{1}{2}U^{\alpha}(10, 10, \ldots) + \frac{1}{2}U^{\alpha}(5, 5, \ldots)\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \\
&\overset{(8)}{=} \left[10^{\rho} + \beta\left(\frac{1}{2}\left(10\left[\frac{1}{1-\beta}\right]^{\frac{1}{\rho}}\right)^{\alpha} + \frac{1}{2}\left(5\left[\frac{1}{1-\beta}\right]^{\frac{1}{\rho}}\right)^{\alpha}\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}}.
\end{aligned}$$

Now, plugging these two results in (13) and (14) makes it possible to determine Robinson's decision for any given set of parameter values. As foreshadowed above, we here want to parameterize three situations. Therefore, we set $\beta = 0.9$ and $\rho = -1$ and consider three different degrees of risk aversion implied by $\alpha_1 = -\frac{1}{2}$, $\alpha_2 = -1$ and $\alpha_3 = -2$. This respectively yields

$$U(d;\alpha) = \begin{cases} 0.7284, & \text{if } \alpha = \alpha_1 \\ 0.7117, & \text{if } \alpha = \alpha_2 \\ 0.6819, & \text{if } \alpha = \alpha_3 \end{cases}$$

and

$$U(\hat{d};\alpha) = \begin{cases} 0.7298, & \text{if } \alpha = \alpha_1 \\ 0.7117, & \text{if } \alpha = \alpha_2 \\ 0.6799, & \text{if } \alpha = \alpha_3. \end{cases}$$

Note that, as generally stated above, Robinson prefers earlier over later resolution of uncertainty, i.e. d over $\tilde{d}$, in the case of $\rho \geq \alpha$. Analogously, preference for later resolution is calibrated via $\rho \leq \alpha$. Eventually, $\rho = \alpha$ parametrizes indifference towards the timing of uncertainty resolution (giving rise to the standard model).

4 Solving a basic DSGE model with EZ/KP utility

This section is intended to illustrate the implementation of the theoretical considerations so far into standard applied macroeconomic analysis, which we understand as the approximate solution of the intertemporal decision problem characterizing some model economy by means of value and policy functions and their usage in simulating artificial data to be contrasted with the stylized facts describing the real economy. Since simulation is a computational exercise independent of the EZ specification, our presentation is only concerned with the approximate solution of EZ economies.

For this purpose, we once more come back to a Robinson Crusoe decision problem. This time, we embed it into a more complete model economy in that we describe the interdependence of current consumption and future consumption opportunities by the means of a savings equation while his output is again subject to an exogenous stochastic influence. Formally, Robinson's situation is described as a stochastic control system. Given an initial capital stock, the way how Robinson chooses his consumption path thereby induces a probability tree. The utility of the latter can thus be found through the representation described in section 3 so that the decision problem is stated in terms of utility maximization. Thereafter, we briefly describe how to apply the perturbation methodology of Schmitt-Grohe and Uribe (2004) to approximate this control system's solution. Eventually, we actually compute a second order perturbation for a given parametrization and discuss the most prominent implications of the EZ/KP framework for such applied work.

4.1 Representative agent environment

Robinson uses capital to produce a final good. He can either consume the final good or use it as investment in the capital stock. His planning horizon is infinite,[27] his personal utility stems solely from consumption and his preferences over uncertain consumption paths have an EZ/KP representation as in (12). To keep the notation parsimonious, we avoid time indices wherever possible.

Now, for each $t \in \mathbb{N}$, we consider the state space

$$X := \mathbb{R}_+ \times \mathbb{R}.$$

An element $x = \begin{pmatrix} x_1 & x_2 \end{pmatrix}^T \in X$ denotes a tuple of capital stock x_1 and productivity level x_2. Further, for every $t \in \mathbb{N}$ Robinson chooses from the control space

$$Y := \mathbb{R}_+.$$

The control $y \in Y$ is interpreted as his consumption level. We endow X and Y with their respective standard topologies.

Moreover, let the triplet $(\Omega, \mathcal{A}, P)$ denote a probability space with a stochastic process $\{\epsilon_t\}_{t=1}^{\infty}$ satisfying $\epsilon_t \sim \text{iidN}(0,1)$ for all $t \in \mathbb{N}$. The control system is now determined by its dynamic

$$\begin{array}{rccc} f: & X \times Y \times \mathbb{R} & \to & \mathbb{R}^2 \\ & (x, y, \epsilon) & \mapsto & \begin{pmatrix} e^{x_2} x_1^{\eta} + (1-\delta)x_1 - y \\ \lambda x_2 + \sigma\epsilon \end{pmatrix}, \end{array} \tag{15}$$

where $\eta, \lambda \in (0,1), \delta \in [0,1]$ and $\sigma \geq 0$. Note that f is continuous and thus measurable. The idea behind this construction is that, given a state x, a choice of the control variable y and a stochastic influence ϵ, next period's state is determined through the mapping f. Thereby,

$$e^{x_2} x_1^{\eta}$$

denotes the produced output and δ determines material wear. Thus, the first component of f says that next period's capital stock amounts to output plus the not worn part of this period's capital stock less chosen consumption. In order to make sure that next period's capital stock is nonnegative, the chosen consumption level has to satisfy

$$y \leq e^{x_2} x_1^{\alpha} + (1-\delta)x_1.$$

[27] Apparently, we cannot simply refer to a bequest motive in order to motivate this assumption so we rather interpret this as an approximation of Robinson's actual situation, which is just as much characterized by his knowledge about the finiteness of his horizon as of the lack of knowledge about the exact duration (maybe complemented with his hope for (or fear of) a long life).

We accordingly define for every $x \in X$

$$Y(x) := \{y \in Y | y \leq e^{x_2} x_1^\alpha + (1-\delta)x_1\}$$

as the feasible control space.[28] On the other hand, f's second component says that the productivity level x_2 evolves as a stationary AR(1) process, i.e. next period's productivity additively depends on its current level and an exogenous stochastic iid influence. While λ determines the persistence of the productivity process, σ is the standard deviation of its stochastic influence and therefore scales the uncertainty in our economy.

We next define a strategy as a measurable mapping

$$h : X \to Y,$$

where X and Y are each endowed with their respective Borel σ-algebra. In other words, a strategy assigns a control value $h(x) \in Y$ to every state $x \in X$, i.e. the agent chooses $h(x)$ whenever confronted with the state x. Note that our definitions of state space and strategy do not allow for the consideration of the state or consumption history. We further call a strategy admissible if it results in a feasible control choice. Accordingly, define the space of admissible strategies as

$$\Pi := \{h\colon\ X \to Y | h \text{ measurable}, h(x) \in Y(x),\ \text{f.a. } x \in X\}.$$

Now, given such a discrete time stochastic control system, an initial state $x_0 \in X$ and an admissible strategy $h \in \Pi$, we recursively define the system's solution under x_0 and h as the state process $\left\{X_t^{x_0,h}\right\}_{t=0}^{\infty}$ and the control process $\left\{Y_t^{x_0,h}\right\}_{t=0}^{\infty}$, where

$$\begin{aligned} X_0^{x_0,h} &:= x_0, \\ Y_t^{x_0,h} &:= h(X_t^{x_0,h}), \text{ f.a. } t \in \mathbb{N}, \\ X_{t+1}^{x_0,h} &:= f(X_t^{x_0,h}, Y_t^{x_0,h}, \epsilon_{t+1}), \text{ f.a. } t \in \mathbb{N}. \end{aligned}$$

Note first that $X_t^{x_0,h}$ and $Y_t^{x_0,h}$ are well-defined random variables for all $t \in \mathbb{N}$ because of the measurability of f, h and ϵ_t, for all $t \in \mathbb{N}$. Second, the solution's recursive construction further reveals that, for all $t \in \mathbb{N}$, $X_t^{x_0,h}$ and

[28] The basic terminology mostly follows Kreps and Porteus (1979).

$Y_t^{x_0,h}$ are already measurable in the random variables $\epsilon_1, \ldots, \epsilon_t$. I.e., if we denote by

$$\mathcal{F}_t := \sigma(\epsilon_1, \ldots, \epsilon_t), t \in \mathbb{N},$$

the σ-algebra generated by $\epsilon_1, \ldots, \epsilon_t$, with $\mathcal{F}_0 := \{\emptyset, \Omega\}$, and further by $\mathcal{F} := \{\mathcal{F}_t\}_{t=0}^{\infty}$ the filtration generated by $\{\epsilon_t\}_{t=0}^{\infty}$, we find the solution $\left\{X_t^{x_0,h}\right\}_{t=0}^{\infty}$ and $\left\{Y_t^{x_0,h}\right\}_{t=0}^{\infty}$ to be $\mathcal{F}$-adapted. Thus, our modeling reasonably assumes Robinson to base his control choice only on already observed ϵ-values. Thereby, the realization of ϵ_t becomes observable for him at the beginning of period t. Moreover, because of our AR(1)/iid assumption it holds for all $s > t$ that given $X_t^{x_0,h}$, the random state $X_s^{x_0,h}$ is independent of $\mathcal{F}_t$. Hence, the solution $\left\{X_t^{x_0,h}\right\}_{t=0}^{\infty}$ and therefore $\left\{Y_t^{x_0,h}\right\}_{t=0}^{\infty}$ both have the Markov property.

4.2 Induced temporal lotteries

Having laid out the basic framework and notation, we next want to describe how to assign a temporal lottery to a tuple of initial state and admissible strategy. Therefore, arbitrarily fix $x_0 \in X$ and $h \in \Pi$. To find the probability tree in D that is induced by this tuple, we consider the solution of the stochastic control system for the control process under x_0 and h. We proceed inductively as we treat the solution in the t^{th} step as if, regarding the uncertainty of future realizations as of period t, we were only interested in their joint distribution.

In the first step, we thus restrict attention with regard to the solution of the control process $\left\{Y_t^{x_0,h}\right\}_{t=0}^{\infty}$ only to $Y_0^{x_0,h} = h(x_0)$ in $t = 0$ and the induced joint probability distribution over $\mathbb{R}_+^{\infty}$ from period 1 on. More precisely, define the stochastic process

$$Y^{x_0,h} \colon \Omega \to \mathbb{R}_+^{\infty}, \omega \mapsto (Y_1^{x_0,h}(\omega), Y_2^{x_0,h}(\omega), \ldots).$$

Next, set $m_1^{x_0,h}$ as the induced image measure over $\mathcal{B}\left(\mathbb{R}_+^{\infty}\right)$, i.e.

$$m_1^{x_0,h} := P \circ (Y^{x_0,h})^{-1},$$

such that for all $B \in \mathcal{B}\left(\mathbb{R}_+^{\infty}\right)$ it holds that

$$\begin{aligned} m_1^{x_0,h}(B) &= P\left(\{\omega \in \Omega | Y^{x_0,h}(\omega) \in B\}\right) \\ &= P\left(\{\omega \in \Omega | (Y_1^{x_0,h}(\omega), Y_2^{x_0,h}(\omega), \ldots) \in B\}\right). \end{aligned}$$

This way, we find the mapping

$$\begin{array}{rrcl} \iota_1: & X \times \Pi & \to & D_1 \\ & (x_0, h) & \mapsto & \left(Y_0^{x_0,h}, m_1^{x_0,h}\right) = \left(h(x_0, h), m_1^{x_0,h}\right). \end{array}$$

We now pursue inductively. Suppose the desired mappings $\iota_1, \ldots, \iota_t$ have already been constructed and we now want to also consider the structure of the induced probability tree until period t. We therefore define for all $B \in \mathcal{B}_t$

$$\begin{aligned} m_{t+1}^{x_0,h}(B) &:= P(\{\omega \in \Omega | \iota_t(X_1^{x_0,h}(\omega), h) \in B\}) \\ &= P \circ \iota_t(X_1^{x_0,h}(\cdot), h)^{-1}(B) \\ &= P_{\iota_t(X_1^{x_0,h}, h)}(B). \end{aligned}$$

The idea behind this definition is the following. Next period's state under x_0 and h is the random variable $X_1^{x_0,h}$. The random t-stage probability tree that is induced by $X_1^{x_0,h}$ and h is given by $\iota_t\left(X_1^{x_0,h}, h\right)$. Thus, the probability distribution over such trees is given by the image measure that is induced by $\iota_t\left(X_1^{x_0,h}(\cdot), h\right)$. Consequently, we define

$$\begin{array}{rrcl} \iota_{t+1}: & X \times \Pi & \to & D_{t+1} \\ & (x_0, h) & \mapsto & \left(Y_0^{x_0,h}, m_{t+1}^{x_0,h}\right) = \left(h(x_0, h), m_{t+1}^{x_0,h}\right). \end{array}$$

Finally, we set

$$\begin{array}{rrcl} \iota: & X \times \Pi & \to & D \\ & (x_0, h) & \mapsto & (\iota_1(x_0, h), \iota_2(x_0, h), \ldots). \end{array}$$

Consistency In order for the mapping $\iota(\cdot, \cdot)$ to be well-defined, the elements of the image sequence have to be consistent in the sense introduced in subsection 2.3.2. To put it another way, if the previous construction was carried out correctly, it has to hold for all $t \in \mathbb{N}$ that

$$\iota_t(x_0, h) = g_t\left(\iota_{t+1}(x_0, h)\right).$$

Since on the one hand

$$g_t\left(\iota_{t+1}(x_0, h)\right) = g_t\left(h(x_0), m_{t+1}^{x_0,h}\right) = \left(h(x_0), f_t(m_{t+1}^{x_0,h})\right)$$

and on the other hand

$$\iota_t(x_0, h) = \left(h(x_0), m_t^{x_0,h}\right),$$

it remains to be shown for all $t \in \mathbb{N}$ that

$$f_t(m_{t+1}^{x_0,h}) = m_t^{x_0,h}.$$

We prove this by induction in appendix A.1. Thus, via ι we can naturally assign the corresponding temporal lottery in D to a tuple of initial state and admissible strategy.

Unique measure over induced trees From subsection 2.3 we further know that we can identify every tree in D with a tuple of current consumption and a probability measure over trees emanating next period via a homeomorphism

$$\Phi : D \to \mathbb{R}_+ \times \mathcal{M}(D).$$

I.e., it holds that $\Phi(\iota(x_0, h)) = (h(x_0), m)$, where $m \in \mathcal{M}(D)$ is the unique measure, which satisfies for all $t \in \mathbb{N}$ and $B \in \mathcal{B}_t$

$$m(\pi_t^{-1}(B)) = m_{t+1}^{x_0,h}(B).$$

We eventually want to show that it holds that

$$\Phi(\iota(x_0, h)) = (h(x_0), P \circ \iota(X_1^{x_0,h}(\cdot), h)^{-1}),$$

i.e.

$$P \circ \iota(X_1^{x_0,h}(\cdot), h)^{-1}(\pi_t^{-1}(B)) = m_{t+1}^{x_0,h}(B).$$

This readily follows from the fact that for all $B \in \mathcal{B}_t$

$$\begin{aligned}
P \circ \iota(X_1^{x_0,h}(\cdot), h)^{-1}(\pi_t^{-1}(B)) = \\
&= P\left(\left\{\omega \in \Omega | \iota(X_1^{x_0,h}(\omega), h) \in \pi_t^{-1}(B)\right\}\right) \\
&= P\left(\left\{\omega \in \Omega | \pi_t(\iota(X_1^{x_0,h}(\omega), h)) \in B\right\}\right) \\
&= P\left(\left\{\omega \in \Omega | \iota_t(X_1^{x_0,h}(\omega), h) \in B\right\}\right) \\
&= m_{t+1}^{x_0,h}(B).
\end{aligned}$$

4.3 Consumption choice

We are now ready to describe Robinson's decision making in this economy. Therefore, define

$$\hat{\iota} := \Phi \circ \iota,$$

i.e. the mapping

$$\begin{array}{rccc} \hat{\iota}: & X \times \Pi & \to & Y \times \mathcal{M}(D) \\ & (x_0, h) & \mapsto & (h(x_0), P \circ \iota(X_1^{x_0,h}(\cdot), h)^{-1}). \end{array}$$

This allows us to indirectly assign a corresponding utility level to the pair (x_0, h). For this purpose, let $U\colon D \to \mathbb{R}_+$ denote a solution to the recursive functional equation

$$U(\Phi^{-1}(c_0, m)) = \left[c_0^\rho + \beta(\mathbb{E}_m U^\alpha)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}}, \beta \in (0,1), 0 \neq \alpha < 1, 0 \neq \rho < 1.$$

as introduced in section 3. We then define the utility mapping

$$\begin{array}{rccc} \hat{U}: & X \times \Pi & \to & \mathbb{R} \\ & (x_0, h) & \mapsto & U(\iota(x_0, h)). \end{array}$$

Using the fact that U is a solution to the recursive equation above, we find

$$\begin{aligned} \hat{U}(x_0, h) &= U(\iota(x_0, h)) = U(\Phi^{-1}(\hat{\iota}(x_0, h))) = U(\Phi^{-1}(h(x_0), P \circ \iota(X_1^{x_0,h}(\cdot), h)^{-1})) \\ &= \left[h(x_0)^\rho + \beta\left(\mathbb{E}_{P \circ \iota(X_1^{x_0,h}(\cdot),h)^{-1}}\left[U^\alpha\right]\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \\ &= \left[h(x_0)^\rho + \beta\left(\mathbb{E}_P\left[U^\alpha(\iota(X_1^{x_0,h}, h))\right]\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \\ &= \left[h(x_0)^\rho + \beta\left(\mathbb{E}_P\left[\hat{U}^\alpha(X_1^{x_0,h}, h)\right]\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \\ &= \left[h(x_0)^\rho + \beta\left(\mathbb{E}_P\left[\hat{U}^\alpha(f(x_0, h(x_0), \epsilon_1), h)\right]\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}}. \end{aligned}$$

Hence, $\hat{U}$ is itself a solution to the recursive functional equation

$$\hat{U}(x, h) = \left[h(x)^\rho + \beta\left(\mathbb{E}_P\left[\hat{U}^\alpha(f(x, h(x), \epsilon), h)\right]\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}}, \tag{16}$$

with $\epsilon \sim N(0,1)$. I.e.,

$$\hat{U}(x,h) = \left[h(x)^\rho + \beta \left(\int_\Omega \hat{U}^\alpha(f(x,h(x),\epsilon(\tilde{\omega})),h)dP(\tilde{\omega}) \right)^{\frac{\rho}{\alpha}} \right]^{\frac{1}{\rho}}$$
$$= \left[h(x)^\rho + \beta \left(\int_\mathbb{R} \hat{U}^\alpha(f(x,h(x),\tilde{\epsilon}),h)\phi(\tilde{\epsilon})d\tilde{\epsilon} \right)^{\frac{\rho}{\alpha}} \right]^{\frac{1}{\rho}},$$

where $\phi(\cdot)$ denotes the standard normal density function.

4.3.1 Decision problem

Having shown how to assign a utility level to a pair of initial state and admissible strategy, we are now able to formulate Robinson's choice problem. Given an initial state x_0, his objective is to find a strategy that maximizes the utility associated with it. I.e. formally Robinson has to solve

$$\max_{h \in \Pi} \hat{U}(x_0,h), \text{ given } x_0.$$

An admissible strategy $h^* \in \Pi$ is thereby called an optimal policy, if it satisfies

$$\hat{U}(x,h^*) \geq \hat{U}(x,h) \text{ f.a. } x \in X, h \in \Pi.$$

Further, we write

$$V(x) := \sup_{h \in \Pi} \hat{U}(x,h)$$

for this consumption problem's value function. Assuming the existence of both, a maximum value, given some $x \in X$, and an optimal policy, we note that the recursive formulation (16) of Robinson's problem directly lends itself to the application of dynamic programming. Thus, from Bellman's Principle of Optimality it follows that an optimal consumption policy has to comply with

$$h^*(x) = \arg\max_{y \in Y(x)} \left[y^\rho + \beta(E_P[V^\alpha(f(x,y,\epsilon))])^{\frac{\rho}{\alpha}} \right]^{\frac{1}{\rho}}, \epsilon \sim N(0,1), \qquad (17)$$

and that the value function must satisfy the intertemporal relation demanded by the (generalized) Bellman equation

$$V(x) = \max_{y \in Y(x)} \left[y^\rho + \beta(E_P[V^\alpha(f(x,y,\epsilon))])^{\frac{\rho}{\alpha}} \right]^{\frac{1}{\rho}}, \epsilon \sim N(0,1). \qquad (18)$$

Necessary optimality conditions In order to actually find an optimal policy, we next use (17) and (18) to derive conditions, which h^* and the induced optimal dynamic have to meet necessarily. We thereby assume the value function to be differentiable. First, an optimal consumption policy must satisfy the Euler equation

$$
\begin{aligned}
0 = \mathbb{E}_P \Bigg[\beta \left(\frac{V^\alpha(f(x, h^*(x), \epsilon))}{\mathbb{E}_P[V^\alpha(f(x, h^*(x), \epsilon))]} \right)^{1-\frac{\rho}{\alpha}} \left(\frac{h^*(f(x, h^*(x), \epsilon))}{h^*(x)} \right)^{\rho-1} \cdot \\
\cdot \left(\eta e^{\lambda x_2 + \sigma\epsilon} (e^{x_2} x_1^\eta + (1-\delta)x_1 - h^*(x))^{\eta-1} + (1-\delta) \right) - 1 \Bigg] .
\end{aligned}
\tag{19}
$$

We derive this result in appendix A.2.

Additionally, in order to clearly distinguish between what can be controlled by Robinson and what is left to pure chance, we introduce further notation. Therefore, denote the deterministic part of the system's dynamic that is induced by following the optimal policy by

$$
f^*(x) := \begin{pmatrix} e^{x_2} x_1^\eta - h^*(x) + (1-\delta)x_1 \\ \lambda x_2 \end{pmatrix} .
$$

Accordingly, we write

$$
f(x, h^*(x), \epsilon) = f^*(x) + \Sigma\epsilon,
$$

with

$$
\Sigma := \begin{pmatrix} 0 \\ \sigma \end{pmatrix}
$$

to denote the optimal dynamic.

Summing up, we have the following conditions, which have to be satisfied by an optimal policy $h^* \in \Pi$, the value function, and the resulting deterministic part f^* of the dynamic

$$
\begin{aligned}
& V(x) - \left[h^*(x)^\rho + \beta \left(\mathbb{E}_P[V(f_1^*(x), f_2^*(x) + \sigma\epsilon)^\alpha] \right)^{\frac{\rho}{\alpha}} \right]^{\frac{1}{\rho}} = 0 \\
& \mathbb{E}_P \Bigg[\beta \left(\frac{V^\alpha(f_1^*(x), f_2^*(x) + \sigma\epsilon)}{\mathbb{E}_P[V^\alpha(f_1^*(x), f_2^*(x) + \sigma\epsilon)]} \right)^{1-\frac{\rho}{\alpha}} \left(\frac{h^*(f_1^*(x), f_2^*(x) + \sigma\epsilon)}{h^*(x)} \right)^{\rho-1} \cdot \\
& \qquad \cdot \left(\eta e^{f_2^*(x) + \sigma\epsilon} (f_1^*(x))^{\eta-1} + (1-\delta) \right) - 1 \Bigg] = 0 \\
& f_1^*(x) - e^{x_2} x_1^\eta + h^*(x) - (1-\delta)x_1 = 0 \\
& f_2^*(x) - \lambda x_2 = 0.
\end{aligned}
\tag{20}
$$

It is generally not possible to solve for the functions h^*, V, and f^* analytically. In the next subsection, we therefore describe a popular method of finding approximations for these functions from this system of equations. Note that neither uniqueness nor existence is ensured by the above set of conditions. This would e.g. demand us to add a transversality condition. The approach taken here, however, is more direct in that it imposes a stability restriction directly on our approximate solution.[29] Further note that the actual feasibility of the approximate solution is typically checked ex post in the simulation results.

4.4 Perturbation

This subsection is intended to illustrate the application of the perturbation approach to EZ environments. Now, due to the fact that the non-linear time aggregation in the EZ representation translates itself directly into the generalized Bellman equation, the DSGE economy studied here has to be slightly modified to match the class of models studied by Schmitt-Grohe and Uribe (2004) (SGU). In particular, by defining an auxiliary variable for Robinson's expected evaluation of continuation utility as of next period, an EZ economy can be fitted into their required structure.

Thus, the following subsection outlines the application of second order perturbation to an EZ economy along the lines of Schmitt-Grohe and Uribe (2004). Thereby, it is detailed enough in order to be self contained and to serve as a complement to the analysis of Caldara et al. (2012), who document on the appropriateness of the perturbation approach to EZ economies.

4.4.1 Method

The method of perturbation relies on the approximation of the optimal policy h^*, the resulting deterministic part of the dynamic f^* and the value function V by means of Taylor polynomials. Thereby, the variance parameter σ in the dynamic f is understood as variable. To make the dependence of the optimal policy, the induced deterministic part of the dynamic and the value function from σ explicit, σ is considered as an additional argument of those functions. I.e. we write

$$h^*(x,\sigma),\ f^*(x,\sigma) \text{ and } V(x,\sigma).$$

[29] See next subsection on its implementation in numerical work.

The Taylor polynomials are thereby expanded around a deterministic fixpoint of the dynamic, i.e. a point $(x, \sigma) = (x_{ss}, 0)$, with $x_{ss} \in X$ satisfying

$$f^*(x_{ss}, 0) = x_{ss}.$$

We call this point the (deterministic) steady state. Now, a simulated solution of the state and control variables along the computed Taylor approximations that results from perturbing the control system from its steady state by allowing $\sigma \neq 0$ is appropriately called a perturbation. For this example, we are computing second order approximations and are hence executing a second order perturbation.

For feasibility, we thus assume h^*, f^* and V to be continuously differentiable in x and σ up to second order. In order to be able to apply the SGU methodology, we further define the auxiliary function

$$W(x, \sigma) := \mathbb{E}_P \left[V((f_1^*(x, \sigma), f_2^*(x, \sigma) + \sigma\epsilon), \sigma)^\alpha \right],$$

for the expectation of the value function to the power of α at next period's random state.

Next, we use the necessary optimality conditions derived above to define the functions $F_i : X \times \mathbb{R} \to \mathbb{R}, i = 1, \ldots, 5$ by

$$\begin{aligned}
F_1(x, \sigma) &:= V(x, \sigma) + \\
&\quad - \left[h^*(x, \sigma)^\rho + \beta \left(W(x, \sigma) \right)^{\frac{\rho}{\alpha}} \right]^{\frac{1}{\rho}} \\
F_2(x, \sigma) &:= \mathbb{E}_P \left[\beta \left(\frac{V^\alpha(f_1^*(x, \sigma), f_2^*(x) + \sigma\epsilon)}{W(x, \sigma)} \right)^{1 - \frac{\rho}{\alpha}} \right. \\
&\qquad \left(\frac{h^*(f_1^*(x, \sigma), f_2^*(x, \sigma) + \sigma\epsilon)}{h^*(x, \sigma)} \right)^{\rho - 1} \\
&\qquad \left. (\eta e^{f_2^*(x) + \sigma\epsilon} f_1^*(x)^{\eta - 1} + (1 - \delta)) \right] - 1 \\
F_3(x, \sigma) &:= W(x, \sigma) - \mathbb{E}_P \left[V((f_1^*(x, \sigma), f_2^*(x, \sigma) + \sigma\epsilon), \sigma)^\alpha \right] \\
F_4(x, \sigma) &:= f_1^*(x, \sigma) - e^{x_2} x_1^\eta + h^*(x, \sigma) - (1 - \delta) x_1 \\
F_5(x, \sigma) &:= f_2^*(x, \sigma) - \rho x_2.
\end{aligned} \tag{21}$$

We thus know that

$$F_i(x, \sigma) \equiv 0, \text{ for all } i = 1, \ldots, 5.$$

Evaluating $F_i(x_{ss}, 0) = 0$ for $i = 1, \dots, 5$ and imposing the above condition for a fixpoint then yields

$$\begin{aligned}
&V(x_{ss}, 0) = \left(\frac{1}{1-\beta}\right)^{\frac{1}{\rho}} h(x_{ss}, 0)\\
&\beta(\eta e^{x_{ss,2}} x_{ss,1}^{\eta-1} + 1 - \delta) = 1\\
&W(x_{ss}, 0) - V(x_{ss}, 0)^{\alpha} = 0\\
&x_{ss,1} - e^{x_{ss,2}} x_{ss,1}^{\eta} + h^*(x_{ss}, 0) - (1-\delta)x_{ss,1} = 0\\
&x_{ss,2} - \rho x_{ss,2} = 0.
\end{aligned}$$

From the last equation it follows that

$$x_{ss,2} = 0.$$

Hence, the second equation delivers Robinson's steady state capital stock

$$x_{ss,1} = \left(\frac{1-\beta(1-\delta)}{\beta\eta}\right)^{\frac{1}{\eta-1}}.$$

The point of expansion for the Taylor polynomials therefore is $(x_{ss}, 0)$ with

$$x_{ss} = \begin{pmatrix} \left(\frac{1-\beta(1-\delta)}{\beta\eta}\right)^{\frac{1}{\eta-1}} \\ 0 \end{pmatrix}.$$

Further, we use the remaining three equations to calculate Robinson's steady state consumption

$$h^*(x_{ss}, 0) = x_{ss,1}^{\eta} - \delta x_{ss,1}$$

and the steady state value of the value function plus the auxiliary function,

$$V(x_{ss}, 0) = \left(\frac{1}{1-\beta}\right)^{\frac{1}{\rho}} h^*(x_{ss}, 0),$$

$$W(x_{ss}, 0) = V(x_{ss}, 0)^{\alpha}.$$

In order to find the Taylor approximations of the optimal policy, the induced deterministic part of the dynamic and the value function around the steady state, we have to compute the derivatives of h^*, f^*, V und W in $(x_{ss}, 0)$ with respect to x_1, x_2 and σ at the steady state.

Now, from $F_i(x, \sigma) \equiv 0, i = 1, \ldots, 5$, it follows that all partial derivatives must be zero, too. I.e. especially at the deterministic steady state it holds that

$$\begin{aligned}
\frac{F_i}{\partial x_1}(x_{ss}, 0) &= 0, i = 1, \ldots, 5, \\
\frac{F_i}{\partial x_2}(x_{ss}, 0) &= 0, i = 1, \ldots, 5, \\
\frac{F_i}{\partial \sigma}(x_{ss}, 0) &= 0, i = 1, \ldots, 5.
\end{aligned} \tag{22}$$

By plugging in the values for $h^*(x_{ss}, 0)$, $f_1^*(x_{ss}, 0)$, $f_2^*(x_{ss}, 0)$, $V(x_{ss}, 0)$ and $W(x_{ss}, 0)$ derived above, (22) is a system of 15 polynomial equations of at most second order in 15 unknowns, namely the partial derivatives of h^*, f_1^*, f_2^*, V and W with respect to x_1, x_2 and σ at the steady state. Solving this system of equations thus yields the sought for first derivatives. In order to be able to pin down the polynomial coefficients uniquely, we additionally demand the dynamic to be stable, or equivalently demand its Jacobian, with respect to x_1, x_2 and σ, evaluated at the steady state, to only have eigenvalues with modulus less than unity.

Next, to find the second order derivatives, we accordingly compute the second order derivatives of all F_i, $i = 1, \ldots, 5$, at the steady state with respect to x_1, x_2 and σ and additionally plug in the already calculated values of the first order derivatives. This yields a (now linear) system of equations in the unknown second order derivatives of h^*, f_1^*, f_2^*, V and W with respect to x_1, x_2 and σ at the steady state. Its necessarily unique solution completes the required computations for a second order perturbation.

4.4.2 Why at least second order?

In this subsection, we want to briefly summarize why it is sensible to at least perform second order approximations of EZ economies. This is a consequence of the certainty equivalence property of first order perturbations as proved in Schmitt-Grohe and Uribe (2004). It states that the coefficients of a linear approximation are independent of the degree of uncertainty in the economy. This is already problematic in general. Most prominently, the expectation of a linearly perturbed variable turns out to equal its deterministic steady state value, entirely independent of σ. Thus, in terms of artificial data generated by a linear approximation of our model, simulating ergodic return time series from pseudorandom iidN(0,1) shocks will yield vanishing simulated risk premia on average, independent of the assumed degree of risk aversion.

Moreover, as demonstrated by van Binsbergen et al. (2012), first order perturbation coefficients are also independent of the risk aversion parameter α. Thus, generalizing a model towards the EZ/KP class leads to identical results as with standard EU preferences ($\alpha = \rho$) if both variants feature the same EIS.[30]

Third, (relative) welfare cost measures of business cycle volatility are typically based on "risky steady state" comparisons of the approximated value function, i.e. on

$$V(x_{ss}, \sigma).$$

Such evaluations thus also demand at least second order approximation.

4.4.3 Numerical example

In this subsection, we use an actual parametric example of our model's second order perturbation to demonstrate where the assumption of nonindifference towards the timing of uncertainty resolution explicitly impacts applied work. Additionally, we begin with pointing at the different ways through which the two parameters of primary interest, EIS and α, affect such second order perturbations.

All figures in this subsection display three different parameterizations of Robinson's attitude towards the timing of uncertainty resolution. Thereby, green lines denote a scenario in which Robinson prefers later resolution, i.e. $\text{EIS}^{-1} > 1 - \alpha$, red lines denote the indifference scenario $\text{EIS}^{-1} = 1 - \alpha$, and black lines display an early resolution case $\text{EIS}^{-1} < 1 - \alpha$. Besides, we fix a quarterly calibration $\beta = 0.95, \eta = 0.36, \lambda = 0.9, \sigma = 0.0072, \delta = 0.011$.

Key parameters First, figure 10 compares the respective effects of α and EIS on the computed approximation of the optimal consumption policy h^*, displayed at the steady state value of productivity as a function of capital only. The left graph results from the calibrations EIS $= 0.5$ with $\alpha = \{-4, -1, 0.5\}$ to calibrate late, indifference and early resolution preferences, respectively. Similarly, the right graph displays the resulting approx-

[30] Note that while this renders the calibration of α ineffective, the EZ/KP representation (at least) still allows α to be set independently of the EIS.

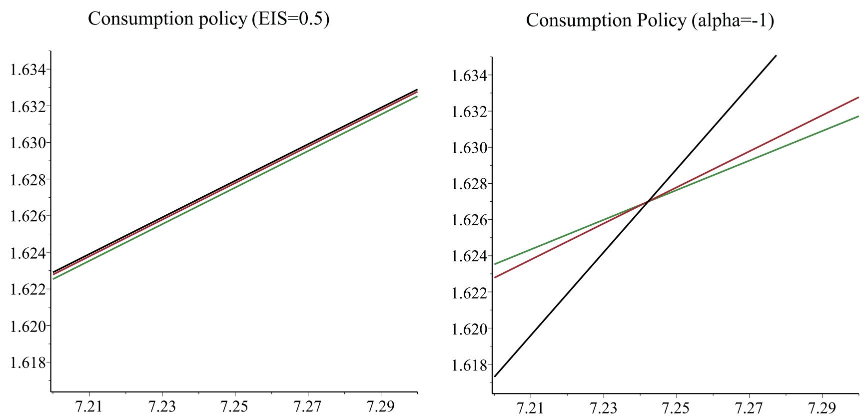

Figure 10: Key parameters (1/2)

imation for $\alpha = -1$ with EIS $= \{0.25, 0.5, 5\}$. It shows that while the risk aversion parameter only shifts the consumption policy, the EIS also exerts influence on its slope. This is generally true. Moreover, scaling uncertainty via σ also only affects the policy's ordinate intercept.[31]

Figure 11 offers another perspective on each parameter's impact on the solution displaying the consumption policy's response to a once only positive shock in ϵ of magnitude 1. The studied scenarios are identical to the figure above. It shows how it is less the attitude towards uncertainty resolution but much rather the EIS directly that impacts the response of macroeconomic quantities. Note how the nature of the effect of EIS is evident in the right graph. The smaller his EIS, the more Robinson strives for a smooth consumption path. Accordingly is a smaller EIS (and therefore rather late resolution preferences for some fixed calibration of α) reflected in larger persistence in the consumption's response.[32]

Nonindifference Another crucial variable in empirical macroeconomics is the equity premium. We finish this discussion with some remarks on its replicability. Therefore, note that Robinson's risky next period return on

[31] See Schmitt-Grohe and Uribe (2004) and van Binsbergen et al. (2012) on this limited effect of α on second order perturbations.

[32] Note the differences in the scenarios' convergence levels, i.e. their *risky* steady state.

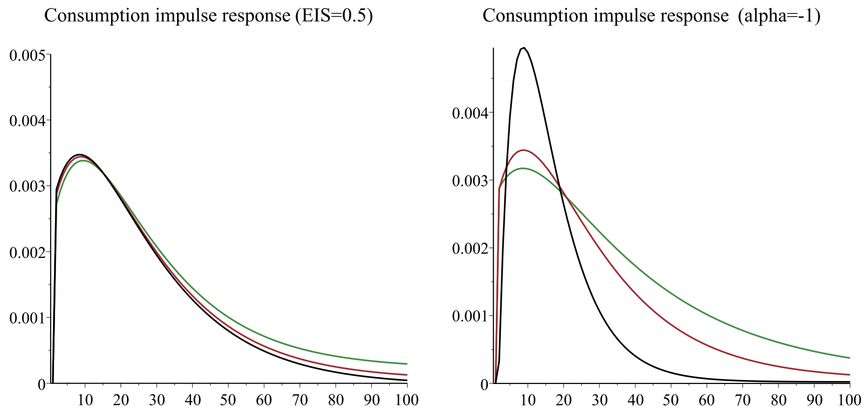

Figure 11: Key parameters (2/2)

equity (RoE) is given by

$$\text{RoE} := \eta e^{\lambda x_2 + \sigma\epsilon}(e^{x_2} x_1^\eta + (1-\delta)x_1 - h^*(x))^{\eta-1} + (1-\delta).$$

Consequently, Robinson's Euler condition (19) has the interpretation of a Lucas (1978) equation such that his stochastic discount factor (SDF) is

$$\text{SDF} := \beta \left(\frac{V^\alpha(f_1^*(x,\sigma), f_2^*(x) + \sigma\epsilon)}{\mathbb{E}_P[V^\alpha(f_1^*(x,\sigma), f_2^*(x) + \sigma\epsilon)]} \right)^{1-\frac{\rho}{\alpha}} \left(\frac{h^*(f_1^*(x,\sigma), f_2^*(x,\sigma) + \sigma\epsilon)}{h^*(x)} \right)^{\rho-1}.$$

Accordingly, if there was a risk free asset available to Robinson, its next period return (r^f) would have to satisfy

$$1 = \mathbb{E}_P [\text{SDF}]\, r^f.$$

Therefore, his expected premium for bearing equity risk

$$EP := \mathbb{E}_P \left[\text{RoE} - r^f\right]$$

would be of magnitude

$$EP = -r^f \text{Cov}[\text{SDF}, \text{RoE}]$$

such that the covariance between SDF and RoE is found to be its key driver.

It is evident from the primitive perturbation system (21) that an assumed nonindifference with respect to the timing of uncertainty resolution enters

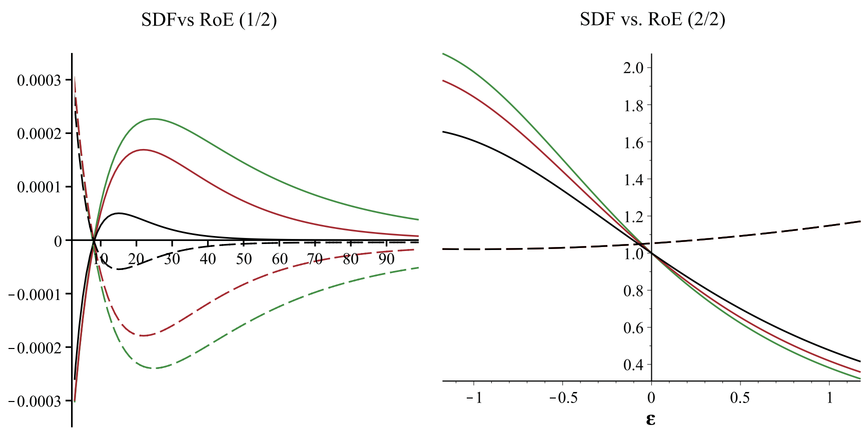

Figure 12: Nonindifference

Robinson's solution through the SDF. Figure 12 therefore illustrates how different preference scenarios affect the SDF and the covariance between SDF and RoE. Both its graphs are generated from a second order perturbation at $\alpha = -1$ and display the SDF (solid lines) and the RoE (dashed lines) in the three parametric scenarios from above, i.e. EIS $= \{0.25, 0.5, 5\}$.

The left graph displays impulse responses to an ϵ shock as above. It first again shows how the convergence to respective *risky* steady state levels is slower for later resolution calibrations. It secondly also shows a more pronounced countercyclicity of SDF and RoE indicating a larger negative covariance. This is further confirmed by the right hand side graph. It explicitly plots both SDF and RoE as functions in (possible realizations of) ϵ with capital and technology at their respective deterministic steady state values. While the RoE is largely unaffected, the SDF shows a stronger negative comovement for smaller EIS (later resolution). This shows how the EZ representation offers applied macroeconomists a channel for the replication of the empirical equity premium. One may impose a strong enough aversion to nonsmooth consumption without having to set the risk aversion parameter unreasonably high.

5 Maple-Matlab toolbox

As outlined above, for perturbation it is necessary to find the derivatives of the functions that constitute our model's equilibrium, all evaluated at the steady state, up to the desired order of the Taylor polynomials. As suggested by Judd (1996) in his introductory remarks, we delegate such computation to a computer algebra system. This section provides some noteworthy details about the associated files, which can be downloaded from the authors homepages.[33]

5.1 Overview

We provide a Maple-based toolbox for perturbation of DSGE models. The core mw-file has an intuitive structure that fits a wide variety of model economies. The systems of equations that determine the perturbation coefficients are derived analytically. Hence, in comparison to purely numerical perturbation packages, our solution is more precise and importantly allows for an intuitive and easy implementation of approximations up to any desired order.[34] In fact, the provided procedures already allow for third order perturbation but the code is straightforwardly extended to higher orders.

The quadratic system of equations for the first order coefficients can be solved either through a general (analytical or a numerical) Maple-internal nonlinear solver or using the generalized Schur decomposition implemented in Matlab.[35]

Alongside, our toolbox features a number of test devices to check for the quality of the solution and provides all necessary information about the solution in the form of txt-output. In order for these files to be correctly stored, the user must create a folder named output in the same directory where the core mw-file is located.

[33] Christopher Heiberger: www.wiwi.uni-augsburg.de/vwl/maussner/lehrstuhl/heiberger_en.html or Halvor Ruf: http://www.wiwi.uni-augsburg.de/vwl/maussner/lehrstuhl/ruf_en.html.

[34] Of course, the practicability of higher order perturbation is nevertheless restricted by computation time.

[35] This solution method makes use of the Matlab-link provided by Maple.

5.2 Brief documentation

Functionality The core `mw`-file requires the user to enter the set of equilibrium conditions that defines the DSGE model under consideration. This is done conveniently using the worksheet's `Math` mode. Some variable x is entered with the suffix 1 if associated with the current period ($x1$) and with suffix 2 for next period ($x2$). The toolbox thereafter demands the user to list the equation numbers (`fktn`), the endogenous states (`xname`), exogenous states (`zname`), the control variables (`uname`), and the shocks (`shocks`). Static equations within the set of equilibrium conditions that are easily solved for (e.g. definitional equations) can be listed as auxiliary functions (`hilfsf`) alongside the associated auxiliary variables (`hilfsvar`) to facilitate the nonlinear solving procedures. Next, the user is asked to enter the model's calibration (`parameter`) and finally its deterministic steady state solution (`ss`) in terms of parameter names (not values). The remaining steps are automatized and briefly outlined in the following paragraphs.

The toolbox's core procedure is `getlsg_fneu`. Using the information entered as described above, it solves for the Taylor polynomials coefficients up to third order (`thrd`) using the chosen solution method for the first order system (`mode`), where `mode` $= 1$ selects the analytical solver, `mode` $= 4$ the numerical solver, and `mode` $= \{2, 3\}$ respectively executes a generalized Schur decomposition using the `Maple-Matlab` link.`getlsg_fneu` itself calls six subprocedures to be sketched below.

First, `transf` transforms the conveniently entered equilibrium conditions into the perturbation logic, i.e. control variables and next period's states as functions of the current states and the perturbation parameter, subject to the model's shocks. Next, `getss` computes the deterministic steady state and writes it into the `steadystate.txt` file.

The subseqent step depends on the chosen solution method. For `mode` $= \{1, 4\}$, the solve routine calls `getgls`. This subprocedure in turn calls `glsys` which generates the basic systems of equations by differentiating the equilibrium conditions and calling `ew` to compute the expected value of all equations. The latter is done analytically making use of the assumed (mutual) independence and standard normality of the shocks. In particular, it iteratively factors the equations as polynomials in the respective shocks and then multiplies the resulting coefficients by these shocks' moments which are determined by the double factorial formula. Thereafter, the equations are returned at the calibrated parameter values. For `mode` $= \{2, 3\}$, `getmatrix` is called. This subprocedure again uses the system of equations generated by `glsys` but now makes further use of its formal structure. In particular, it generates the matrix pencil A, B as in Schmitt-Grohe and Uribe (2009) and writes it into the two respectively named `txt` files.

In the remaining steps, `getlsg_fneu` computes the sought for derivatives of the optimal policy and dynamic by the chosen method and collects them in the arrays $J = \begin{bmatrix} J_1 & J_2 & J_3 \end{bmatrix}$ and $H = \begin{bmatrix} H_1 & H_2 & H_3 \end{bmatrix}$. The three elements of J are the Jacobians of the endogenous dynamic (J_1), the exogenous dynamic (J_2), and the optimal policy (J_3), while the three elements of H are the associated cubes consisting of the respective Hessian matrices in the same order. These Jacobians and Hessian cubes are finally written into accordingly named `txt` files. Note that in the Schur-based method ($\texttt{mode} = \{2, 3\}$), the uniqueness of the quadratic first order system's solution is attained by construction. For $\texttt{mode} = \{1, 4\}$, uniqueness is forced through additionally imposing the dynamic's Jacobian to only have eigenvalues within the unit circle.

Quality of solution In order to check and document the quality of the computed solution, the toolbox features the following instruments. First, the deterministic steady state values are inserted into the primitive equilibrium conditions and the equations numerical deviation from 0 is printed. Second, for all solution methods, `getlsg_fneu` checks the uniqueness of the first order solution and aborts the computation with an error message in case of multiple equilibria. Third, we also provide the maximum norms of the differences $A - Q^T AZ^T$ and $B - Q^T BZ^T$, respectively, to check whether the computed matrices Q and Z actually constitute a Schur decomposition. Fourth, the QZ decomposition can be executed in both orders, A, B ($\texttt{mode} = 2$) and B, A ($\texttt{mode} = 3$), such that a cross check may support confidence in the solution. This, of course, holds just as much for crosschecking over the other available solution methods.

6 Conclusion

In this paper, we summarized the crucial elements of the EZ representation, demonstrated the application of the Schmitt-Grohe and Uribe (2004) approach, and provide a flexible computer algebra toolbox for its application. As an immediate implication for applied work, we find the EIS to play a more prominent role than the risk aversion parameter α and (consequently) suspect the *late* resolution case to be likely to give rise to higher equity premia.

In light of the latter results, we would like to close with two remarks. First, the incomplete disentanglement of comparative risk aversion from EIS, which gives rise to nonindifference with respect to resolution timing, might itself be undesirable.[36] A more general representation in which risk aversion and EIS are completely separated would not only provide more flexibility for applied work but also help to clarify the roles played by each of these key aspects of preferences.[37] Second, the parameter of atemporal consumption risk aversion might simply be an inadequate measure of risk aversion in intertemporal decision problems.[38]

[36] Note that it is not only difficult to rationalize why preferences for later resolution generate larger premia, it is even less evident how to correctly calibrate nonindifference.

[37] See Weil (1990), p. 33, for a conjecture on how to achieve further disentanglement. Note however, that there might also be an "inherent inseparability" as suspected by Epstein and Zin (1989), p. 953.

[38] See Träger (2011) and Swanson (2012) for related theoretical work.

A Appendix

A.1 Consistency of induced trees

We want to prove by induction that the induced trees are indeed consistent. Therefore, we additionally introduce further notation. Define for any $\tau \in \mathbb{N}$ the solution of the state process $\{X_t^{x_0,h,\tau}\}_{t=0}^{\infty}$ and the control process $\{Y_t^{x_0,h,\tau}\}_{t=0}^{\infty}$ under the initial state x_0 and the admissible strategy h, that arises, if we shift the stochastic process $\{\epsilon_t\}_{t=1}^{\infty}$ by τ periods, i.e.

$$\begin{aligned}
X_0^{x_0,h,\tau} &:= x_0 \\
Y_t^{x_0,h,\tau} &:= h(X_t^{x_0,h,\tau}), \text{ f.a. } t \in \mathbb{N}, \\
X_{t+1}^{x_0,h,\tau} &:= f(X_t^{x_0,h,\tau}, Y_t^{x_0,h,\tau}, \epsilon_{\tau+t+1}), \text{ f.a. } t \in \mathbb{N}.
\end{aligned}$$

Note that because of the iid assumption of the stochastic process $\{\epsilon_t\}_{t=0}^{\infty}$, the probability distribution for this process is the same as for the non-shifted solution. Yet, the specific realizations for some arbitrary $\omega \in \Omega$ may differ.

Now, starting with $t = 1$, we–according to subsection 2.3.2–find for all $B \in \mathcal{B}\left(\mathbb{R}_+^{\infty}\right)$

$$\begin{aligned}
f_1(m_2^{x_0,h})(B) &= \int_{\mathbb{R}_+ \times \mathcal{M}(\mathbb{R}_+^{\infty})} \tilde{m}_1 \left(\left\{y \in \mathbb{R}_+^{\infty} | (\tilde{c}_1, y) \in B\right\}\right) dm_2^{x_0,h}(\tilde{c}_1, \tilde{m}_1) \\
&= \int_{\mathbb{R}_+ \times \mathcal{M}(\mathbb{R}_+^{\infty})} \tilde{m}_1 \left(\left\{y \in \mathbb{R}_+^{\infty} | (\tilde{c}_1, y) \in B\right\}\right) dP \circ \left(\iota_1(X_1^{x_0,h}(\cdot), h)\right)^{-1} (\tilde{c}_1, \tilde{m}_1) \\
&= \int_{\Omega} m_1^{X_1^{x_0,h}(\tilde{\omega}),h} \left(\left\{y \in \mathbb{R}_+^{\infty} | (Y_0^{X_1^{x_0,h}(\tilde{\omega}),h}, y) \in B\right\}\right) dP(\tilde{\omega}).
\end{aligned}$$

Now, note that for an arbitrary but fixed $\tilde{\omega} \in \Omega$ it holds for all $B \in \mathcal{B}\left(\mathbb{R}_+^{\infty}\right)$

that

$$m_1^{X_1^{x_0,h}(\tilde{\omega}),h}\left(\left\{y \in \mathbb{R}_+^\infty | (Y_0^{X_1^{x_0,h}(\tilde{\omega}),h}, y) \in B\right\}\right) =$$
$$= P\left(\left\{\omega \in \Omega | (Y_1^{X_1^{x_0,h}(\tilde{\omega}),h}(\omega), Y_2^{X_1^{x_0,h}(\tilde{\omega}),h}(\omega), \dots) \in \left\{y \in \mathbb{R}_+^\infty | (Y_0^{X_1^{x_0,h}(\tilde{\omega}),h}, y) \in B\right\}\right\}\right)$$
$$= P\left(\left\{\omega \in \Omega | (Y_0^{X_1^{x_0,h}(\tilde{\omega}),h}(\omega), Y_1^{X_1^{x_0,h}(\tilde{\omega}),h}(\omega), Y_2^{X_1^{x_0,h}(\tilde{\omega}),h}(\omega), \dots) \in B\right\}\right)$$
$$= P\left(\left\{\omega \in \Omega | (Y_0^{X_1^{x_0,h}(\tilde{\omega}),h,1}(\omega), Y_1^{X_1^{x_0,h}(\tilde{\omega}),h,1}(\omega), \dots) \in B\right\}\right)$$
$$= P\left(\left\{\omega \in \Omega | (Y_0^{X_1^{x_0,h}(\tilde{\omega}),h,1}(\omega), Y_1^{X_1^{x_0,h}(\tilde{\omega}),h,1}(\omega), \dots) \in B\right\} \middle| \left\{\omega \in \Omega | X_1^{x_0,h}(\omega) = X_1^{x_0,h}(\tilde{\omega})\right\}\right)$$
$$= P\left(\left\{\omega \in \Omega | (Y_1^{x_0,h}(\omega), Y_2^{x_0,h}(\omega), \dots) \in B\right\} \middle| \left\{\omega \in \Omega | X_1^{x_0,h}(\omega) = X_1^{x_0,h}(\tilde{\omega})\right\}\right).$$

Thereby, the third last equality makes use of the above definition of a solution that is shifted by $\tau = 1$ steps. The second last equality then follows from the fact that in this shifted solution there is no ϵ_1 such that it is stochastically independent of the condition. Thereby, the expressions in the last two lines denote the conditional probability for the respective random sequence given the event $X_1^{x_0,h} = X_1^{x_0,h}(\hat{\omega})$. Because of the shifting, the random sequences in the last two lines are identical. The base case now follows from the properties of conditional expectations. To see this, we define

$$C := \left\{\omega \in \Omega | (Y_1^{x_0,h}(\omega), Y_2^{x_0,h}(\omega), \dots) \in B\right\}$$

and find

$$f_1(m_2^{x_0,h})(B) =$$
$$= \int_\Omega P\left(\left\{\omega \in \Omega | (Y_1^{x_0,h}(\omega), Y_2^{x_0,h}(\omega), \dots) \in B\right\} \middle| \left\{\omega \in \Omega | X_1^{x_0,h}(\omega) = X_1^{x_0,h}(\tilde{\omega})\right\}\right) dP(\tilde{\omega})$$
$$= \int_\Omega P\left(C \middle| \left\{\omega \in \Omega | X_1^{x_0,h}(\omega) = X_1^{x_0,h}(\tilde{\omega})\right\}\right) dP(\tilde{\omega})$$
$$= \int_\Omega \mathbb{E}\left[\mathbb{1}_C \middle| X_1^{x_0,h} = X_1^{x_0,h}(\tilde{\omega})\right] dP(\tilde{\omega}) = \int_\Omega \mathbb{E}\left[\mathbb{1}_C \middle| X_1^{x_0,h}\right](\tilde{\omega}) dP(\tilde{\omega})$$
$$= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_C \middle| X_1^{x_0,h}\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_C | \mathcal{F}_1\right]\right] = \mathbb{E}\left[\mathbb{1}_\Omega \mathbb{E}\left[\mathbb{1}_C | \mathcal{F}_1\right]\right] \overset{\Omega \in \mathcal{F}_1}{=} \mathbb{E}\left[\mathbb{1}_\Omega \mathbb{1}_C\right] = P(C)$$
$$= P\left(\left\{\omega \in \Omega | (Y_1^{x_0,h}(\omega), Y_2^{x_0,h}(\omega), \dots) \in B\right\}\right) = m_1^{x_0,h}(B).$$

Thus, the induction hypothesis (IH) reads

$$f_t(m_{t+1}^{x_0,h}) = m_t^{x_0,h},$$

for some $t \in \mathbb{N}$. Recalling the constructed f_t and g_t from subsection 2.3.2, the inductive step follows for all $t \geq 1$ and $B \in \mathcal{B}_t$, i.e.

$$\begin{aligned}
f_{t+1}(m_{t+2}^{x_0,h})(B) =& m_{t+2}^{x_0,h}(g_t^{-1}(B)) \\
=& P\left(\left\{\omega \in \Omega | \iota_{t+1}(X_1^{x_0,h}(\omega), h) \in g_t^{-1}(B)\right\}\right) \\
=& P\left(\left\{\omega \in \Omega | g_t\left(\iota_{t+1}(X_1^{x_0,h}(\omega), h)\right) \in B\right\}\right) \\
=& P\left(\left\{\omega \in \Omega | g_t\left(Y_1^{x_0,h}(\omega), m_{t+1}^{X_1^{x_0,h}(\omega)}\right) \in B\right\}\right) \\
=& P\left(\left\{\omega \in \Omega | (Y_1^{x_0,h}(\omega), f_t(m_{t+1}^{X_1^{x_0,h}(\omega)})) \in B\right\}\right) \\
\overset{\text{(IH)}}{=}& P\left(\left\{\omega \in \Omega | (Y_1^{x_0,h}(\omega), m_t^{X_1^{x_0,h}(\omega)}) \in B\right\}\right) \\
=& m_{t+1}^{x_0,h}(B).
\end{aligned}$$

This shows that the sequence $(\iota_1(x_0, h), \iota_2(x_0, h), \ldots)$ is consistent and therefore lies in D.

A.2 Euler equation

Differentiating the bracketed right hand side expression of (17) with respect to y, we find the necessary condition for a maximum

$$\begin{aligned}
0 =& \frac{1}{\rho}[\ldots]^{\frac{1}{\rho}-1}\Big(\rho y^{\rho-1} + \\
&+ \beta\frac{\rho}{\alpha}(\mathbb{E}_P[V^\alpha(f(x,y,\epsilon))])^{\frac{\rho}{\alpha}-1}\mathbb{E}_P[\alpha V^{\alpha-1}(f(x,y,\epsilon_t))\frac{\partial V}{\partial x_1}(f(x,y,\epsilon_t))\frac{\partial f_1}{\partial y}(x,y,\epsilon)]\Big) \\
=& V^{1-\rho}(x)\Big(y^{\rho-1} - \beta(\mathbb{E}_P[V^\alpha(f(x,y,\epsilon))])^{\frac{\rho}{\alpha}-1}\mathbb{E}_P[V^{\alpha-1}(f(x,y,\epsilon))\frac{\partial V}{\partial x_1}(f(x,y,\epsilon_t))]\Big) \\
\Leftrightarrow& y^{\rho-1} - \beta(\mathbb{E}_P[V^\alpha(f(x,y,\epsilon))])^{\frac{\rho}{\alpha}-1}\mathbb{E}_P[V^{\alpha-1}(f(x,y,\epsilon))\frac{\partial V}{\partial x_1}(f(x,y,\epsilon))] = 0.
\end{aligned}$$

Introducing the abbreviation

$$X^{(1)} := f(x, h^*(x), \epsilon),$$

h^* must hence satisfy

$$h^*(x)^{\rho-1} - \beta(\mathbb{E}_P[V^\alpha(X^{(1)})])^{\frac{\rho}{\alpha}-1}\mathbb{E}_P[V^{\alpha-1}(X^{(1)})\frac{\partial V}{\partial x_1}(X^{(1)})] = 0. \tag{23}$$

Further, by the envelope theorem, we find the derivative of the value function with respect to x_1 from (18)

$$\begin{aligned}\frac{\partial V}{\partial x_1}(x) &= \frac{1}{\rho}[\ldots]^{\frac{1}{\rho}-1}\beta\frac{\rho}{\alpha}\mathbb{E}_P[V^\alpha(f(x,h^*(x),\epsilon)])^{\frac{\rho}{\alpha}-1}\cdot \\ &\quad\cdot\mathbb{E}_P\left[\alpha V^{\alpha-1}(f(x,h^*(x),\epsilon)\frac{\partial V}{\partial x_1}(f(x,h^*(x),\epsilon)\frac{\partial f_1}{\partial x_1}(x,h^*(x),\epsilon)\right] \\ &= V^{1-\rho}(x)\beta\mathbb{E}_P[V^\alpha(f(x,h^*(x),\epsilon))])^{\frac{\rho}{\alpha}-1}\cdot \\ &\quad\cdot\mathbb{E}_P\left[V^{\alpha-1}(f(x,h^*(x),\epsilon))\frac{\partial V}{\partial x_1}(f(x,h^*(x),\epsilon))(\eta e^{x_2}x_1^{\eta-1}+(1-\delta))\right],\end{aligned}$$

or again abbreviated

$$\frac{\partial V}{\partial x_1}(x) = V^{1-\rho}(x)(\eta e^{x_2}x_1^{\eta-1}+(1-\delta))\beta\mathbb{E}_P[V^\alpha(X^{(1)})])^{\frac{\rho}{\alpha}-1}\mathbb{E}_P\left[V^{\alpha-1}(X^{(1)})\frac{\partial V}{\partial x_1}(X^{(1)})\right]. \tag{24}$$

From (23) we see

$$\frac{\partial V}{\partial x_1}(x) = V^{1-\rho}(x)(\eta e^{x_2}x_1^{\eta-1}+(1-\delta))\,(h^*(x))^{\rho-1}. \tag{25}$$

Now, iterating this equation forward by one period and plugging it into (23) yields

$$\begin{aligned}(h^*(x))^{\rho-1} &= \beta\mathbb{E}_P[V^\alpha(X^{(1)})])^{\frac{\rho}{\alpha}-1}\cdot \\ &\qquad\cdot\mathbb{E}_P[V^{\alpha-1}(X^{(1)})V^{1-\rho}(X')(\eta e^{X_2^{(1)}}\left(X_1^{(1)}\right)^{\eta-1}+(1-\delta))\left(h^*(X^{(1)})\right)^{\rho-1}],\end{aligned}$$

or equivalently

$$\begin{aligned}0 &= \mathbb{E}_P\left[\beta\mathbb{E}_P[V^\alpha(X^{(1)})])^{\frac{\rho}{\alpha}-1}V^{\alpha-\rho}(X^{(1)})(\eta e^{X_2^{(1)}}\left(X_1^{(1)}\right)^{\eta-1}+(1-\delta))\left(\frac{h^*(X^{(1)})}{h^*(x)}\right)^{\rho-1}-1\right] \\ &= \mathbb{E}_P\left[\beta\left(\frac{V^\alpha(X^{(1)})}{\mathbb{E}_P[V^\alpha(X^{(1)})]}\right)^{1-\frac{\rho}{\alpha}}\left(\frac{h^*(X^{(1)})}{h^*(x)}\right)^{\rho-1}(\eta e^{X_2^{(1)}}\left(X_1^{(1)}\right)^{\eta-1}+(1-\delta))-1\right].\end{aligned}$$

References

Allais, Maurice F. C. 1953. "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine." *Econometrica* 21 (October): 503–546.

Altug, Sumru, and Pamela Labadie. 2008. *Asset Pricing for Dynamic Economies*. Campbridge University Press.

Backus, David K., Bryan R: Routledge, and Stanley E. Zin. 2008. "Recursive Preferences." In *The New Palgrave Dictionary of Economics Online*, edited by Steven N. Durlauf and Lawrence E. Blume, Second Edition. Palgrave Macmillan.

Caldara, Dario, Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez, and Wen Yao. 2012. "Computing DSGE Models with Recursive Preferences and Stochastic Volatility." *Review of Economic Dynamics* 15 (April): 188–206.

Chew, Soo H. 1983. "A Generalization of the Quasilinear Mean with Application to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox." *Econometrica* 51 (July): 1065–1092.

———. 1989. "Axiomatic Utility Theories with the Betweenness Property." *Annals of Operations Research* 19 (1): 273–298 (December).

Chew, Soo H., and Larry G. Epstein. 1989. "The Strucure of Preferences and Attitudes towards the Timing of the Resolution of Uncertainty." *International Economic Review* 30 (1): 103–117 (February).

Debreu, Gerard. 1954. "Representation of a Preference Ordering by a Numerical Function." Chapter XI of *Decision Processes*, edited by Robert M. Thrall, C. H. Coombs, and R. L. Davis, 159–165. Wiley.

Epstein, Larry, Emmanuel Farhi, and Tomasz Strzalecki. 2014. "How Much Would You Pay to Resolve Long-Run Risk?" *American Economic Review*.

Epstein, Larry G. 1992. "Behavior under Risk: Rcent Developments in Theory and Application." Chapter 1 of *Advances in Economic Theory Sixth World Congress*, edited by Jean-Jaques Laffont, Volume II of *Econometric Society Monographs*, 1–63. Campbridge University Press.

Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57 (July): 937–969.

Fisher, Irving. 1930. *The Theory of Interest.* The Macmillan Company. Reprinted by Porcupine Press, 1977.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature* 40 (June): 351–401.

Judd, Kenneth L. 1996. "Approximation, perturbation, and projection methods in economic analysis." Chapter 12 of *Handbook of Computational Economics*, edited by H. M. Amman, D. A. Kendrick, and J. Rust, Volume 1 of *Handbook of Computational Economics*, 509–585. Elsevier.

Koopmans, Tjalling C. 1960. "Stationary Ordinal Utility and Impatience." *Econometrica* 28 (2): 287–309 (April).

Kreps, David M., and Evan L. Porteus. 1978. "Temporal Resolution of Uncertainty and Dynamic Chioce Theory." *Econometrica* 46 (January): 185–200.

———. 1979. "Dynamic Choice Theory and Dynamic Programming." *Econometrica* 47 (1): 91–100 (January).

Lucas, Robert E. Jr. 1978. "Asset Prices in an Exchange Economy." *Econometrica* 46 (November): 1429–1445.

Olson, Mancur, and Martin Bailey. 1981. "Positive Time Preference." *Journal of Political Economy* 89 (February): 1–25.

Parthasarathy, Kalyanapuram R. 1967. *Probability Measures on Metric Spaces.* Academic Press.

Raiffa, Howard. 1970. *Decision Analysis.* Addison-Wesley.

Samuelson, Paul A. 1937. "A Note on Measurement of Utility." *Review of Economic Studies* 4 (February): 155–161.

Schmitt-Grohe, Stephanie, and Martin Uribe. 2004. "Solving Dynamic General Equilibrium Models Using a Second Order Approximation to the Policy Function." *Journal of Economic Dynamics and Control* 28 (January): 755–775.

———. 2009. "Perturbation Methods for the Numerical Analysis of DSGE Models (Lecture Notes)."

Selden, Larry. 1978. "A New Representation of Preferences over "Certain x Uncertain" Consumption Pairs: The "Ordinal Certainty Equivalent" Hypothesis." *Econometrica* 46 (5): 1045–1060 (September).

Spence, Michael, and Richard Zeckhauser. 1972. "The Effect of the Timing of Consumption Decisions and the Resolution of Lotteries on the Choice of Lotteries." *Econometrica* 40 (March): 401–403.

Swanson, Eric T. 2012. "Risk Aversion and the Labor Margin in Dynamic Equilibrium Models." *American Economic Review* 102 (4): 1663–91 (June).

Träger, Christian P. 2011. "Intertemporal Risk Aversion - or - Wouldn't it be Nice to Tell Whether Robinson Crusoe is Risk Averse?" *CUDARE Working Papers*, vol. 1102 (August).

van Binsbergen, Jules H., Jesús Fernández-Villaverde, Ralph S.J. Koijen, and Juan F. Rubio-Ramírez. 2012. "The term structureof interestrates in a DSGE model with recursive preferences." *Journal of Monetary Economics* 59 (7): 634–648 (November).

Weil, Phillipe. 1990. "Nonexpected Utility in Macroeconomics." *The Quarterly Journal of Economics* 105 (1): 29–42 (February).