Heterogeneous Impatience and Dynamic Inconsistency

Chiaki Hara*

Institute of Economic Research, Kyoto University

February 7, 2013

Abstract

In a continuous-time equilibrium model of heterogeneous consumers, we formulate and prove the statement that the more heterogeneous the consumers are in their impatience, the more dynamically consistent the representative consumer is. We also characterize these orderings within some parameterized families of the distribution of the consumers’ discount rates. We apply these results to the term structure of interest rates, and, in particular, to accommodate heterogeneous impatience in the model of Cox, Ingersoll, and Ross (1985).

JEL Classification Codes: D51, D53, D61, D81, D91, E43, G12.

Keywords: Discount factor, discount rate, representative consumer, term structure of interest rates, cumulant-generating function.

---

*I am grateful to the audiences at the University of Bielefeld (February 2011), the twelfth SAET Conference at the University of Queensland (July 2012), the Ajou International Workshop in Financial Economics and Mathematics at Ajou University (July 2012), the University of Turku (August 2012), and the First Asian Quantitative Finance Conference at the National University of Singapore (January 2013). The paper is part of the academic projects on Economic Analysis of Intergenerational Issues: Searching for Further Development, funded by the Grant-in-Aid for Specially Promoted Research (grant number 22000001), and on Frontier of Game theory: Theory and Practice, funded by the Grant-in-Aid (S) (grant number 2023001), both from Japan’s Ministry of Education, Culture, Sports, Science and Technology. My email address is hara.chiaki.7x@kyoto-u.ac.jp.
1 Introduction

Asset transactions are often motivated by the heterogeneity in consumers’ characteristics. More risk averse consumers unload the risks they are faced with and less risk averse ones take them over with premiums. Optimistic consumers invest more in risky assets, while pessimistic consumers invest mostly into riskless bonds. More patient consumers save more to enjoy higher consumptions in the future, and impatient consumers may well borrow to enjoy higher consumption immediately. The raison d’être of asset markets is precisely to cater for diverse needs for asset transactions by heterogeneous consumers.

Heterogeneity in consumers’ characteristics have implications not only on risk and intertemporal allocations but also on asset pricing. The impact on asset pricing can probably be best understood by constructing the representative consumer. The representative consumer is a fictitious consumer whose marginal utility process, evaluated along the aggregate consumption process, is a state price deflator. For example, the representative consumer of individual consumers having utility functions of constant and unequal relative risk aversion has a utility function of strictly decreasing relative risk aversion (Franke, Stapleton, and Subrahmanyam (1999) and Hara, Huang and Kuzmics (2007)); and the representative consumer of individual consumers having constant but unequal subjective discount rates has discount rates that are a strictly decreasing function of time (Weitzman (2001), Gollier and Zeckhauser (2005), and Lengwiler (2005)). The consequences of these are that the derivative assets with convex payoff functions is underestimated if the coefficients of constant relative risk aversion are erroneously assumed to be equal (Franke, Stapleton, and Subrahmanyam (1999) and Hara, Huang, and Kuzmics (2007)); and that the term structure of interest rates is more downward sloping in the case of heterogeneous subjective discount rates than in the case of homogeneous subjective discount rates (Lengwiler (2005)).

While Weitzman (2001), Gollier and Zeckhauser (2005), and Lengwiler (2005) showed that the heterogeneity in individual consumers’ subjective discount rates gives rise to a dynamically inconsistent representative consumer, they did not clarify how the degree of heterogeneity is related to the degree of dynamic inconsistency. In particular, they did not show whether a more heterogeneous economy give rise to a more dynamically inconsistent representative consumer. Without such a result, we would be left unsure whether a homogeneous economy could be
so singular that no inference should be drawn from comparisons with a homogeneous economy to comparisons between two heterogeneous economies.

The purpose of this paper is to give a precise formulation to the statement that the more heterogeneous the subjective discount rates of the individual consumers are, the more dynamically inconsistent the representative consumer is. To do so, we need to give a notion of the “more dynamically inconsistent than” relation and the “more heterogeneous than” relation, between two heterogeneous economies. For the former, we use the notion by Prelec (2004), which was introduced for a single decision maker having a utility function over “single dated outcomes”. For the latter, we introduce, following Jouini and Napp (2007, Lemma 4.1), the notion of an approximate wealth-weighted distribution of subjective discount rates of individual consumers (hereafter referred to, more simply, as the distribution of discount rates). As it is a probability measure on the strictly positive part $R^+\mathbb{R}$ of the real line, we can define its cumulant-generating function on the non-positive part $-R_\mathbb{R}$ of the real line. Then our main result (Theorem 2) shows that the more convex the cumulant-generating function of the distribution of discount rate is, the more dynamically inconsistent the representative consumer is, and the converse also holds.

To understand what the convexity assumption means, recall that the convexity of a twice differentiable function is measured by its curvatures, which are the ratios of the second derivatives to the first derivatives. It is well known that the first and second derivatives of the cumulant-generating function at $s \leq 0$ coincide with the mean and variance of the probability measure of which the density function with respect to the distribution of discount rates is proportional to the exponential function $q \mapsto \exp(sq)$. Hence the curvature of the cumulant-generating function at $s$ is the ratio of the variance to the mean of the probability measure with a density proportional to $q \mapsto \exp(sq)$, which can be regarded as a degree of the heterogeneity of the distribution of subjective discount rates. Our main result, therefore, formalizes the notion that the more heterogeneous the subjective discount rates are, the more dynamically inconsistent the representative consumer is.

When applied to the representative consumer, the notion of dynamic inconsistency should be taken with a pinch of salt. The reason is that the representative consumer is a fictitious agent, who does not autonomously choose asset portfolios but is merely used to determine the state-price density at equilibrium. In partic-
ular, no consumer in the model of this paper wishes to revise his portfolio at any subsequent period, because every individual consumer has a constant discount rate and is, hence, dynamically consistent. The dynamically inconsistency of the representative consumer should, therefore, be regarded as a way to describe the nature of asset prices that cannot emerge in the dynamically consistent representative consumer model.

When it comes to obtaining implications on the term structure of interest rates, it is often unnecessary to establish dynamic consistency. As we will see, the more-dynamically-inconsistent-than relation turns is nothing but the monotone likelihood ratio condition of the two discount rate functions under consideration. The natural weaker notion of comparison is the single-crossing property of these two functions, and, for the analysis of the term structure of interest rates, it is often sufficient to establish the single-crossing property. We will show that it can be derived from the single-crossing property of the derivatives of the two corresponding cumulant-generating functions, and implies the single-crossing property of the forward-rate curves, the yield curves, and the short-rate process. In addition, we generalize the model of the term structure of interest rates by Cox, Ingersoll, and Ross (1985) to accommodate heterogeneous impatience.

If the distributions of discount rates belong to some parametric family, the task of determining whether one such distribution has a more convex cumulant-generating function than another is relatively easy. In Section 7, we give two examples of such families, one consisting of Gamma distributions and the other consisting of Bernoulli distributions. Among other things, we show that the shape parameter of Gamma distributions is irrelevant for the measure of dynamic inconsistency; and that within the family of binomial distributions, the second-order stochastic dominance relation is insufficient for the more-dynamically-inconsistent-than relation.

This paper is organized as follows. The setup and preliminary results are presented in Section 2. The measure of heterogeneity of impatience is introduced in Section 3. In Section 4, we see when the utility weights defining the representative consumer do or do not approximate wealth shares. The more-dynamically-inconsistent-than relation of Prelec (2004) is reviewed in Section 5. The main results are stated in Section 6. Parameterized examples are given in Sections 7. Applications to the term structure of interest rates are explored in Section 8. The results are summarized and a future research topic is suggested in Section 9.
2 Setup and Preliminary Results

2.1 Economy

The economy is subject to uncertainty, which is represented by a probability space \((\Omega, \mathcal{F}, P)\). The time span is \(R_+ = [0, \infty)\), which is of continuous time and infinite length. The gradual information revelation is represented by a filtration \((\mathcal{F}(t))_{t \in R_+}\). There is only one type of good on each time and state.

We allow the number of consumers present in the economy to be finite or infinite. Formally, we let \((A, \mathcal{A}, \nu)\) be a finite measure space of (names of) consumers. If \(A\) is a finite set, \(\mathcal{A}\) is the power set of \(A\), and \(\nu\) is the counting measure on \(A\), then the consumption sector consists of finitely many consumers. If, on the other hand, \(A\) is the unit interval \([0, 1]\), \(\mathcal{A}\) is the Borel \(\sigma\)-field \(\mathcal{B}([0, 1])\), and \(\nu\) is (the restriction of) the Lebesgue measure on \(\mathcal{B}([0, 1])\), then the consumption sector consists of infinitely many consumers, each of whom is negligible in size relative to the total population of the economy. For each \(B \in \mathcal{A}\), \(\nu(B)\) is the proportion of consumers in \(B\) relative to the entire consumption sector.

We assume that the consumers have time-additive expected utility functions over consumption processes, which exhibit constant and equal relative risk aversion, and constant but possibly unequal discount rates. Formally, let \(\gamma > 0\) and define \(u : R_+^+ \to R\) by

\[
\begin{align*}
u(x) = & \begin{cases} 
\ln x & \text{if } \gamma = 1, \\
\frac{x^{\gamma-1} - 1}{\gamma - 1} & \text{otherwise,}
\end{cases} 
\end{align*}
\]

for every \(x \in R_+^+\). Let \(\rho : A \to R_+^+\) be measurable, where \(R_+^+\) is endowed with the Borel \(\sigma\)-field \(\mathcal{B}(R_+^+)\). Then the utility function \(U_a\) of consumer \(a\) over consumption processes is defined by

\[
U_a(c_a) = E \left( \int_0^\infty \exp(-\rho(a)t)u(c_a(t)) \, dt \right),
\]

where \(c_a = (c_a(t))_{t \in R_+}\).\(^1\) Although the assumption of constant and equal relative

---

\(^1\)This and other integrals in the subsequent analysis need not be well defined without additional assumptions on \(c_a\) and other stochastic processes. But the subsequent argument depends only on the first-order conditions of (utility or social welfare) maximization problems, which must necessarily hold whenever there is a solution to the problem under consideration. We shall therefore be implicit about these additional assumptions.
risk aversion is quite stringent, there is a good reason to restrict our attention to this case. In fact, Hara (2009, Corollary 2) showed that if consumers had unequal coefficients of constant relative risk aversion, then the representative consumer’s discount factor function, to be specified below, would not be well defined.

For each $a \in A$, let $e_a$ be the initial endowment process of consumer $a$, which is an $\mathbb{R}_+$-valued adapted process. The aggregate endowment process, $\int_A e_a \, d\nu(a)$, is denoted by $e$.

### 2.2 Arrow-Debreu equilibrium

A state-price deflator is an $\mathbb{R}_+^+$-valued adapted process. The utility maximization problem of consumer $a$ under a state-price deflator $\pi$ is

$$
\max_c \quad U_a(c_a) \\
\text{subject to} \quad E \left( \int_0^\infty \pi(t)(c(t) - e_a(t)) \, dt \right) = 0.
$$

(1)

We say that a state price deflator $\pi$ and an allocation $(c_a)_{a \in A}$ of consumption processes constitute an Arrow-Debreu equilibrium if $\int_A c_a \, d\nu(a) = \int_A e_a \, d\nu(a)$ and for every $a \in A$, $c_a$ is a solution to the utility maximization problem (1) of $a$ under the state-price deflator $\pi$. Although we shall not elaborate on this point, it is well known that the Arrow-Debreu equilibrium allocations coincide with the equilibrium allocations in complete asset markets.

### 2.3 Representative consumer

By Proposition 10.C of Duffie (2001), for each Arrow-Debreu equilibrium, there is a $\lambda : A \to \mathbb{R}_+^+$ such that the solution to the social welfare maximization problem over allocations of the aggregate consumption process $e$,

$$
\max_{(c_a)_{a \in A}} \quad \int_A \lambda(a)U_a(c_a) \, d\nu(a), \\
\text{subject to} \quad \int_A c_a \, d\nu(a) = c.
$$

(2)

coincides with the equilibrium allocation when $c = e$. It was shown in Hara (2008, Section 2) that the value function of this maximization problem, which is the
representative consumer’s utility function, is given by

\[ U(c) = E \left( \int_0^\infty d(t)u(c(t)) \, dt \right). \tag{3} \]

where

\[ d(t) = (h(t))^\gamma. \tag{4} \]

and

\[ h(t) = \int_A (\lambda(a) \exp(-\rho(a)t))^{1/\gamma} \, d\nu(a). \tag{5} \]

Note that the representative consumer too has constant relative risk aversion equal to $\gamma$. The function $d : \mathbb{R}_+ \to \mathbb{R}$ is the representative consumer’s discount factor function. In order for $U$ to be well defined, it is necessary and sufficient that $h(0) < \infty$, because $\exp(-\rho(a)t) \leq 1$ for every $a \in A$ and every $t \in \mathbb{R}_+$. This is equivalent to saying that the function $a \mapsto (\lambda(a))^{1/\gamma}$ is integrable with respect to $\nu$. As we will see, $d$ is an analytic function. Define the representative consumer’s discount rate function $r : \mathbb{R}_+ \to \mathbb{R}_{++}$ by

\[ r(t) = -\frac{d'(t)}{d(t)}. \tag{6} \]

then

\[ \frac{d(t_2)}{d(t_1)} = \exp \left( -\int_{t_1}^{t_2} r(t) \, dt \right) \]

whenever $0 \leq t_1 < t_2$. Thus $r$ represents the representative consumer’s continuously compounded instantaneous subjective discount rate as a function of time. Unlike the case of individual consumers, this is not constant but varies with $t$ unless all individual consumers have the same discount rate.

The equilibrium state price deflator is given by $du'(e) = (d(t)u'(e(t)))_{t \in \mathbb{R}_+}$. Thus the price at time $t_1$, relative to the current consumption, of the discount bond with maturity $t_2 > t_1$ is equal to

\[ E_{t_1} \left( \frac{d(t_2)u'(e_{t_2})}{d(t_1)u'(e_{t_1})} \right) = \frac{d(t_2)}{d(t_1)} E_{t_1} \left( \frac{u'(e_{t_2})}{u'(e_{t_1})} \right) = \exp \left( -\int_{t_1}^{t_2} r(t) \, dt \right) E_{t_2} \left( \frac{u'(e_{t_2})}{u'(e_{t_1})} \right). \]

We are interested in how the heterogeneity in $\rho$ affect these bond prices.
3 Measure of heterogeneous impatience

Recall from (4) and (5) that the representative consumer’s discount factor function \( d \) can be written as

\[
d(t) = \left( \int_A (\lambda(a))^{1/\gamma} \exp \left( -\frac{\rho(a)t}{\gamma} \right) d\nu(a) \right)^\gamma.
\]  

(7)

Define a probability measure \( \mu^1 \) on \((R_+, \mathcal{B}(R_+))\) by letting

\[
\mu^1(B) = \int_{\rho^{-1}(B)} (\lambda(a))^{1/\gamma} d\nu(a)
\]

for every \( B \in \mathcal{B}(R_+) \). Alternatively, \( \mu^1 \) can be defined by first letting \( \nu^1 \) the finite measure on \( A \) for which

\[
\frac{d\nu^1}{d\nu} = \left( \int_A (\lambda(a))^{1/\gamma} d\nu(a) \right)^{-1} \lambda^{1/\gamma}
\]

and then letting \( \mu^2 = \nu^2 \circ \rho^{-1} \). Denote its moment-generating function by \( M^1 \), that is, \( M^1(s) = \int_{R_+} \exp(sq) d\mu^1(q) \). Then

\[
d(t) = \left( \int_A (\lambda(a))^{1/\gamma} d\nu(a) \right)^\gamma \left( M^1 \left( -\frac{t}{\gamma} \right) \right)^\gamma
\]

Denote its cumulant-generating function by \( K^1 \), that is, \( K^1(s) = \ln M^1(s) \). Then

\[
r(t) = (K^1)' \left( -\frac{t}{\gamma} \right).
\]  

(9)

It is well known if the first two moments exist (are finite), then \( K^1 \) is twice differentiable, with \((K^1)'(0)\) equal to the mean of \( \mu^1 \) and \((K^1)''(0)\) equal to the variance of \( \mu^1 \). Denote by \( \mu^1(s) \) the probability measure on \( R_+ \) such that

\[
\frac{d\mu^1(s)}{d\mu^1}(q) = \exp \left( sq - K^1(s) \right)
\]

for every \( q \in R_+ \). Then, by Morris (1982, Section 2), for every \( s \), \((K^1)'(s)\) and \((K^1)''(s)\) are equal to the mean and variance of \( \mu^1(s) \). Thus, \((K^1)'(s) > 0\) for every \( s \), and, unless \( \mu^1 \) is concentrated on a single point, \((K^1)''(s) > 0\) for every \( s \). Hence \((K^1)''(s)/(K^1)'(s)\) is the ratio of the variance to the mean, and can be
considered as a measure of dispersion of the probability measure \( \mu \). We shall take it as the measure of the individual consumers’ discount rates.

4 A Digression on utility weights and wealth shares

As can be seen in (7), the representative consumer’s discount factor function \( d : R_+ \rightarrow R_{++} \) is, roughly, the weighted average of the individual consumers’ counterparts, where the weights are equal to the \( (\lambda(a))^{1/\gamma} \) and the latter have been chosen so that the solution to the social welfare maximization problem (2) is the Arrow-Debreu equilibrium allocation. Hence, the measure of heterogeneous impatience defined in Section 3 and its relation to the measure of dynamic consistency, to be given in Section 6, are stated in terms of the distribution of discount rates weighted by the \( (\lambda(a))^{1/\gamma} \). Unfortunately, this aspect of our results is an impediment to its applications, because the \( (\lambda(a))^{1/\gamma} \) are unobservable. In this section, we investigate whether and how the weighted distribution of discount rates with weights \( (\lambda(a))^{1/\gamma} \) can be approximated by something more likely to be observable, such as the population distribution and the wealth distribution.

Recall that for each consumer \( a \), \( \lambda(a) \) is equal to the reciprocal of the Lagrange multiplier associated with the budget constraint of the utility maximization problem (1). Given the concavity of \( u \), the value of \( (\lambda(a))^{1/\gamma} \) is positively related to his wealth levels. On the other hand, the unweighted average, or, more precisely, the population-weighted average of the individual consumers’ discount factors would be

\[
\left( \int_A \exp \left( -\frac{\rho(a) t}{\gamma} \right) \, d\nu(a) \right)^{\gamma}.
\]

(10)

Not surprisingly, this does not coincide with the discount factor function \( d \), because the wealthier the individual consumer, the more pronounced his impact on the representative consumer’s discount factor, while (10) pays no attention to this effect. Yet, when \( \gamma = 1 \), the population-weighted average (10) has been the focus of the analysis of collective decision making, such as the cost-benefit analysis. In fact, Weitzman (2001) asked more than two thousand PhD-level economists which discount factor should be used to discount the cost and benefit of mitigating climate change, and the answers turned out to follow a gamma distribution with mean 3.96% and standard deviation 2.94%. He then showed that the discount
rate function derived from the population-weighted average (10) is a hyperbolic function of \( t \) whenever the population-weighted distribution of discount rates is a Gamma function.

From the viewpoint of the economic analysis, it would be better if we could write the discount factor function \( d \) in terms of the wealth distribution rather than the weights \( (\lambda(a))^{1/\gamma} \), because we may have access to the data set of income distributions of an economy or experimental results in which subjects’ discount rates are solicited and their incomes can somehow be inferred. Formally, the wealth share function \( \theta : A \to \mathbb{R}^{++} \) is defined by

\[
\theta(a) = \frac{E \left( \int_0^\infty \pi(t)e_a(t) \, dt \right)}{E \left( \int_0^\infty \pi(t)e(t) \, dt \right)},
\]

and the wealth-weighted average of the individual consumers’ discount factors would be

\[
\left( \int_A \theta(a) \exp \left( -\frac{\rho(a)t}{\gamma} \right) \, d\nu(a) \right)^\gamma.
\]

This is, in general, different from (7). Even worse, knowing the wealth share function \( \theta \) does not allow us to identify the utility weight function \( \lambda \) and hence the discount factor function \( d \). We shall not provide any formal result on this point, but a simple example might still be illustrating.

Consider two economies, indexed by \( n = 1, 2 \). They share the same set of consumers, \( A_n = \{1, 2\} \), and the same common coefficient \( \gamma \) of constant relative risk aversion, which is less than one. The two consumers in each economy have different discount rates but these values are the same in the two economies. In symbols, for \( \rho_n : A_n \to \mathbb{R}^{++} \), we let \( \rho_1(1) = \rho_2(1) < \rho_1(2) = \rho_2(2) \). The two economies are different only initial endowments: the aggregate initial endowment processes \( e_n \) are deterministic for both \( n = 1, 2 \), but the ratio between the two, \( e_2/e_1 \), is strictly increasing over time. Each consumer owns half of the aggregate initial endowment at each time and state. Take any equilibrium of each of the two economies. Then the aggregate wealth is shared equally between the two consumers at both equilibria. For each \( n = 1, 2 \), let \( \lambda_n : A_n \to \mathbb{R}^{++} \) be the utility weight function that correspond to the equilibrium of economy \( n \). We then claim that \( \lambda_1(1)/\lambda_2(1) > \lambda_1(2)/\lambda_2(2) \). Hence the representative consumer’s discount factor functions \( d_1 \) and \( d_2 \) are different in the two economies.
The reason for \( \lambda_1(1)/\lambda_1(2) > \lambda_2(1)/\lambda_2(2) \) is intuitive: since the aggregate endowment processes \( e_n \) are deterministic, the equilibrium consumption paths \( c_{an} \), with \( n = 1, 2 \) and \( a \in A_n \), are all deterministic. Moreover, since \( \rho_n(1) < \rho_n(2) \), \( c_{1n}/c_{2n} \) is strictly increasing over time. Yet, since \( e_2/e_1 \) is strictly increasing, if \( \lambda_1(1)/\lambda_1(2) = \lambda_2(1)/\lambda_2(2) \), then consumer 1 would consume more in economy 2 than in economy 1. Let \( \pi_n \) be the equilibrium state-price deflator for economy \( n \), then it is deterministic for each \( n \), \( \pi_2/\pi_1 \) is strictly decreasing because \( e_2/e_1 \) is strictly increasing, but the ratio of the product, \( \pi_2 e_2/\pi_1 e_1 \), is strictly increasing because \( \gamma < 1 \). Hence, if \( \lambda_1(1)/\lambda_1(2) = \lambda_2(1)/\lambda_2(2) \) and if consumer 1’s budget constraint were satisfied in economy 1, then it would be violated in economy 2. Therefore, to have his budget constraint satisfied in both economies, it is necessary that \( \lambda_1(1)/\lambda_1(2) > \lambda_2(1)/\lambda_2(2) \).

Although the powered utility weight function \( \lambda^{1/\gamma} \) uniquely determines the representative consumer’s discount factor function \( d \), the wealth share function \( \theta \) does not. It is, thus, meaningful investigate to what extent \( \lambda^{1/\gamma} \) and \( \theta \) are close to each other. If all the consumers have the same discount rate (that is, \( \rho: A \rightarrow \mathbb{R}^{++} \) is constant valued), the mutual fund theorem holds, in the sense that for the solution \( (c_a)_{a \in A} \) to the maximization problem (2), the fraction \( c_a/c \) is, up to scalar multiplication, equal to \( (\lambda(a))^{1/\gamma} \) at any time in any state. Thus, \( \lambda^{1/\gamma} \) coincides, up to scalar multiplication, with \( \theta \). If consumers do not have the same discount rate, however, the consumption share \( c_a/c \) is not constant across time, and \( \lambda^{1/\gamma} \) is not equal to \( \theta \) even up to scalar multiplication, but Jouini and Napp (2007, Lemma 4.1) gave bounds on the discrepancy of \( \lambda \) from \( \theta \). We shall, therefore, refer to the weighted distribution of the discount rates \( \rho(a) \) with weights \( (\lambda(a))^{1/\gamma} \) as the approximated wealth-weighted distribution of the individual consumers’ discount rates, and explore how the degree of dispersion of this distribution is related to the dynamic inconsistency of the representative consumer.

A weakness of \( \lambda^{1/\gamma} \) is that it is, even up to scalar multiplication, not equal to \( \theta \) even when \( \gamma = 1 \). It turns out, then, that the \( \lambda^{1/\gamma}/\rho \) is, up to scalar multiplication, equal to the \( \theta \). This can be shown as follows.

By the first-order condition and the envelope theorem of the maximization problem (2),

\[
c_a(t) = (\lambda(a) \exp(-\rho(a) t))^{1/\gamma} (\pi(t))^{-1/\gamma} = \frac{(\lambda(a) \exp(-\rho(a) t))^{1/\gamma}}{h(t)} e(t) \tag{13}
\]
\[ \pi(t) = d(t)u'(e(t)) = (h(t))^{\gamma} (e(t))^{-\gamma}. \]

Thus, by the budget constraint and Fubini’s theorem,
\[
E \left( \int_0^\infty \pi(t)e_a(t) \, dt \right) = \int_A E \left( \int_0^\infty \pi(t)e_a(t) \, dt \right) \, d\nu(a)
= \int_0^\infty \left( \int_A \lambda(a) \exp(-\rho(a)t) \right)^{1/\gamma} (h(t))^{\gamma-1} \left( (e(t))^{1-\gamma} \right) \, dt
= \int_0^\infty (h(t))^{\gamma} E \left( (e(t))^{1-\gamma} \right) \, dt.
\]

Hence
\[
E \left( \int_0^\infty \pi(t)e(t) \, dt \right) = \int_A E \left( \int_0^\infty \pi(t)e_a(t) \, dt \right) \, d\nu(a)
= \int_0^\infty \left( \int_A \lambda(a) \exp(-\rho(a)t) \right)^{1/\gamma} \, d\nu(a) \left( h(t) \right)^{\gamma-1} E \left( (e(t))^{1-\gamma} \right) \, dt
= \int_0^\infty (h(t))^{\gamma} E \left( (e(t))^{1-\gamma} \right) \, dt.
\]

Therefore,
\[
\theta(a) = \frac{\int_0^\infty \left( \lambda(a) \exp(-\rho(a)t) \right)^{1/\gamma} (h(t))^{\gamma-1} \left( (e(t))^{1-\gamma} \right) \, dt}{\int_0^\infty (h(t))^{\gamma} E \left( (e(t))^{1-\gamma} \right) \, dt}. \tag{14}
\]

If, in particular, \( \gamma = 1 \), then
\[
\theta(a) = \frac{\int_0^\infty \lambda(a) \exp(-\rho(a)t) \, dt}{\int_0^\infty h(t) \, dt}
= \frac{\int_0^\infty \lambda(a) \exp(-\rho(a)t) \, dt}{\int_0^\infty \lambda(b) \exp(-\rho(b)t) \, dt} \, d\nu(b)
= \frac{\lambda(a)}{\rho(a)} \frac{\lambda(b)}{\rho(b)} \, d\nu(b).
\]

Hence, when \( \gamma = 1 \), \( \lambda/\rho \) is a scalar multiple of \( \theta \).

It is, therefore, reasonable to guess that \( \lambda^{1/\gamma} / \rho \) approximates \( \theta \) better than \( \lambda^{1/\gamma} \) if \( \gamma \) is close to one. If this is indeed the case, our subsequent analysis should really
be stated and proved in terms of the weighted distributions with weights \( \lambda^{1/\gamma} / \rho \) rather than \( \lambda^{1/\gamma} \). But we shall opt for \( \lambda^{1/\gamma} / \rho \) for two reasons. First, as we will see, they admit a clearer relation between the degree of heterogeneity of the individual consumers’ discount rates and the degree of the representative consumer’s dynamic inconsistency. Second, as we will also see, the results in terms of weights \( \lambda^{1/\gamma} / \rho \) can be derived from the corresponding results in terms of weights \( \lambda^{1/\gamma} \).

5 Measure of dynamic inconsistency

Prelec (2004) introduced an at-least-as-decreasingly-impatient-as relation between two utility functions over single dated outcomes, that is, functions of the form \( d(t)u(x) \) defined over consumption levels \( x \) consumed at time \( t \), where \( d : \mathbb{R}_+ \to (0, 1) \) is strictly decreasing, and \( u : \mathbb{R}_+ \to \mathbb{R}_+ \) is strictly increasing and satisfies \( u(0) = 0 \). The following definition is a variant of the relation, is concerned with discount factor functions, rather than utility functions over timed consumptions, and which is equivalent to Prelec’s definition if \( u \) is continuous.

**Definition 1** A discount factor function \( d_1 \) is at-least-as-decreasingly-impatient-as another discount factor function \( d_2 \) if for every \( (t_0, t_1, t_2, \tau) \in \mathbb{R}_+^3 \times \mathbb{R} \),

\[
\frac{d_2(t_0 + t_1)}{d_2(t_0)} \geq \frac{d_2(t_0 + t_1 + t_2 + \tau)}{d_2(t_0 + t_2)}
\]

whenever

\[
\frac{d_1(t_0 + t_1)}{d_1(t_0)} = \frac{d_1(t_0 + t_1 + t_2 + \tau)}{d_1(t_0 + t_2)}.
\]

To understand this definition, first compare the two ratios, \( d_1(t_0 + t_1)/d_1(t_0) \) and \( d_1(t_0 + t_1 + t_2)/d_1(t_0 + t_2) \). The former is the discount factor that \( d_1 \) applies to the time interval \([t_0, t_0 + t_1]\), and the latter is the discount factor that \( d_1 \) applies to the time interval \([t_0 + t_2, t_0 + t_1 + t_2]\). Since they are both applied to time intervals of length \( t_1 \), they would be equal if \( d_1 \) exhibited exponential discounting, that is, if there were a \( q > 0 \) such that \( d_1(t) = \exp(-qt) \) for every \( t \in \mathbb{R}_+ \).\(^2\) However, they can be different, and, more specifically, the former is smaller than the latter if the corresponding discount rate function is decreasing over time, just as in the case of hyperbolic discounting. To compensate the difference between the two ratios, we add an interval of length \( \tau \) (which is positive if the discounting rate function is decreasing).

\(^2\)Then both of the two ratios are equal to \( \exp(-qt_1) \).
decreasing over time, but negative if it is increasing) to the terminal time \( t_0 + t_1 + t_2 \) of the interval, so that the discount rate that \( d_1 \) applies to \([t_0 + t_2, t_0 + t_1 + t_2 + \tau]\) is equal to the discount rate that \( d_1 \) applies to \([t_0, t_0 + t_1]\), as shown in (16). The length \( \tau \) can, therefore, be considered as a measure of decreasing impatience of \( d_1 \). Then (15) states that \( s \) may too large for \( d_2 \), so that the discount factor that \( d_2 \) applies to \([t_0 + t_2, t_0 + t_1 + t_2 + \tau]\) may be smaller than the discount factor that \( d_2 \) applies to the time interval \([t_0, t_0 + t_1]\). In this sense, the impatience of \( d_1 \) decreases at least as rapidly as that of \( d_2 \) as the time interval under consideration is shifted into a more distant future. This is exactly the idea that Definition 1 embodies.

Prelec (2004, Proposition 1) proved the following equivalence on the at-least-as-decreasingly-impatient-as relation.

**Theorem 1 (Prelec (2004))** Let \( d_1 \) and \( d_2 \) be thrice differentiable discount factor functions and \( r_1 \) and \( r_2 \) be the corresponding discount rate functions. Then the following two conditions are equivalent.

1. \( d_1 \) is at least as decreasingly impatient as \( d_2 \).

2. \(-r_1'(t)/r_1(t) \geq -r_2'(t)/r_2(t)\) for every \( t \in R_+ \).

Thanks to this theorem, to determine the ranking of the degree of decreasing impatience between two discount factor functions, it is sufficient to compare the rates of decrease of the corresponding discount rate functions. In the subsequent analysis, we identify how the rate of decrease of the representative consumer’s discount rate function is related to the degree of heterogeneity of individual consumers’ discount rates.

### 6 Comparison between two economies

Consider two economies, \( n = 1, 2 \), each with the space \((A_n, \mathcal{A}_n, \nu_n)\) of (names of) consumers, a discount rate function \( \rho_n : A \rightarrow R_{++} \), and a weighting function \( \lambda_n : A_n \rightarrow R_{++} \). Assume that the consumers of the two economies share the same coefficient \( \gamma \) of constant relative risk aversion. Using \((A_n, \mathcal{A}_n, \nu_n), \rho_n, \) and \( \lambda_n \), define the discount factor function \( d_n : R_+ \rightarrow R_{++} \) in the same way as \( d \) in (4) and (5), and the discount rate function \( r_n : R_+ \rightarrow R_{++} \) in the same way as
Define the probability measure $\mu_n$ in the same way as $\mu^1$ in (8), and let $K^1_n$ be the cumulant-generating function of $\mu^1_n$.

By differentiating both sides of (9), we obtain

$$r'(t) = -\frac{1}{\gamma} K''\left(-\frac{t}{\gamma}\right),$$

(17)

$$-r'(t) = \frac{1}{\gamma} \frac{K''\left(-\frac{t}{\gamma}\right)}{K'(\frac{-t}{\gamma})}.$$  

(18)

The following theorem follows immediately from these equalities.

**Theorem 2** The following two conditions are equivalent.

1. For every $t \geq 0$, $-r'_1(t)/r_1(t) > -r'_2(t)/r_2(t)$.

2. For every $s \leq 0$, $(K^1_1)^''(s)/(K^1_1)'(s) > (K^1_2)^''(s)/(K^1_2)'(s)$.

The first condition of this theorem states that the representative consumer of the first economy is more dynamically inconsistent than the representative consumer of the second economy. Since

$$\frac{d}{dt} \left( \frac{r_1(t)}{r_2(t)} \right) = \frac{r_1(t)}{r_2(t)} \left( \frac{r'_1(t)}{r_1(t)} - \frac{r'_2(t)}{r_2(t)} \right),$$

it is equivalent to the *monotone likelihood ratio condition*, in that $r_1(t)/r_2(t)$ is a strictly decreasing function of $t$. The second condition states that the cumulant-generating function of the weighted distribution $\mu^1_n$ of individual consumers’ discount rates in the first economy is more convex than in the second economy. This condition is equivalent to saying that the variance divided by the mean of the individual consumers’ discount rates is higher in the first economy than in the second whenever their distribution is transformed by a negative exponential density function. The theorem, then, asserts that the representative consumer is more dynamically inconsistent if and only if the cumulant-generating function of the weighted distribution of individual consumers’ discount rates is more convex.

The next theorem deals with weaker conditions, although they are still useful to investigate the term structure of interest rates.

**Theorem 3** The following two conditions are equivalent.
1. For every $t \geq 0$, if $r_1(t) = r_2(t)$, then $r_1'(t) < r_2'(t)$.

2. For every $s \leq 0$, if $(K_1^n)'(s) = (K_2^n)'(s)$, then $(K_1^n)''(s) > (K_2^n)''(s)$.

The condition in the first part of this theorem implies the single-crossing property, in that $r_1$ crosses $r_2$ at most once from above: if $r_1(t_0) = r_2(t_0)$, then $r_1(t) < r_2(t)$ for every $t > t_0$, and $r_1(t) > r_2(t)$ for every $t < t_0$. That is, the discount rate in the first economy is higher than in the second up to a time, after which the former is lower. The second part is the single-crossing property of the $(K_1^n)'$, where $(K_1^n)'$ crosses $(K_2^n)'$ at most once from below. This condition is equivalent to saying that if the weighted distributions of the individual consumers’ discount rates are transformed by a negative exponential density function so that the means with respect to the transformed distribution are equal in the two economies, then the variance is higher in the first economy than in the second. This theorem follows directly from (9) and (17).

As we mentioned toward the end of Section 4, $\lambda^{1/\gamma}/\rho$ seems to approximate the wealth share function $\theta$ than $\lambda^{1/\theta}$. In concluding this section, we touch on a necessary and sufficient condition for a more dynamically inconsistent representative consumer in terms of the weighted distribution induced by $\lambda^{1/\gamma}/\rho$.

Define a probability measure $\mu^2$ on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ by letting

$$\mu^2(B) = \int_{\rho^{-1}(B)} (\lambda(a))^{1/\gamma} \rho(a) d\nu(a)$$

for every $B \in \mathcal{B}(\mathbb{R}^+)$. Alternatively, $\mu^2$ can be defined by first letting $\nu^2$ the finite measure on $A$ for which

$$\frac{d\nu^2}{d\nu} = \left( \int_A (\lambda(a))^{1/\gamma} \rho(a) d\nu(a) \right)^{-1} \frac{(\lambda)^{1/\gamma}}{\rho}$$

and then letting $\mu^2 = \nu^2 \circ \rho^{-1}$. Denote its moment-generating function by $M^2$, that is, $M^2(s) = \int_{\mathbb{R}^+} \exp(sq) d\mu^2(q)$. Since

$$\frac{d\nu^1}{d\nu^2}(a) = \frac{d\nu^1(a)}{d\nu^2(a)} = \psi(a),$$

16
where
\[ \psi = \left( \int_A (\lambda(a))^{1/\gamma} \, d\nu(a) \right)^{-1} \left( \int_A (\lambda(a))^{1/\gamma} \, \rho(a) \, d\nu(a) \right), \]
it can be shown that
\[ \frac{d\mu_1}{d\mu_2}(q) = \frac{d(\nu_1 \circ \rho^{-1})}{d(\nu_2 \circ \rho^{-1})}(q) = \psi q. \tag{20} \]
Thus,
\[ (M^2)'(s) = \int_{R_{++}} \exp(sq) \, d\mu_2(q) = \psi^{-1} \int_{R_{++}} \exp(sq) \, d\mu_1(q) = \psi^{-1} M^1(s). \tag{21} \]
Hence, all the results in terms of \( M^1 \) and \( K^1 \) can be restated in terms of \( M^2 \). In particular, since \( K^1(s) = \ln(M^2)'(s) + \ln \psi \), (6) implies that
\[ r(t) = \frac{(M^2)'' \left( -\frac{t}{\gamma} \right)}{(M^2)' \left( -\frac{t}{\gamma} \right)}. \]
Thus, Theorems 2 and 3 can be restated in terms of the \( M^2_n \).

7 Comparison within a parametrized family

In many applications, we do not compare two particular weighted distributions \( \mu_1 \) and \( \mu_2 \) of individual consumers’ discount rates. Rather, we consider a parameterized family of weighted distributions of individual consumers’ discount rates, say \((\mu^1(\cdot, \alpha, \beta))_{(\alpha, \beta) \in D}\) where \( D \) is an open subset of \( \mathbb{R}^2 \), and determine how the parameter values \((\alpha, \beta)\) are related to the measure of the representative consumer’s dynamic inconsistency. In this section, we give two such families, one consisting of Gamma distributions and the other consisting of Bernoulli distributions.

7.1 Gamma distributions

First, we consider the family of gamma distributions, each with parameters \( \alpha \) and \( \beta \), that is, its density function (with respect to Lebesgue measure) is given by
\[ q \mapsto \frac{\beta^\alpha q^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta q). \tag{22} \]
Gamma distributions are most commonly used as distributions of subjective time
discount rates, as in Weitzman (2001), Gollier and Zeckhauser (2005), and Hara
(2007).

Define \( d(\cdot, \alpha, \beta) \) as the discount factor function when the weighted distribution
\( \mu^1 \) defined in (8) has the density function (22). That is,

\[
d(t, \alpha, \beta) = \left( \int_0^{\infty} \exp\left( -\frac{qt}{\gamma} \right) \frac{\beta^\alpha q^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta q) \, dq \right)^\gamma.
\]

Then let \( r(\cdot, \alpha, \beta) \) be the discount factor function corresponding to \( d(\cdot, \alpha, \beta) \), that is,

\[
r(t, \alpha, \beta) = -\frac{\partial d(t, \alpha, \beta)}{\partial t} \cdot d(t, \alpha, \beta).
\]

The corresponding cumulant-generating function \( K^1(\cdot, \alpha, \beta) \) is given by

\[
K^1(s, \alpha, \beta) = \alpha(\ln \beta - \ln(\beta - s)).
\]

Thus

\[
\frac{\partial K^1}{\partial s}(s, \alpha, \beta) = \frac{\alpha}{\beta - s},
\]

(23)

\[
\frac{\partial^2 K^1}{\partial s^2}(s, \alpha, \beta) = \frac{1}{\beta - s}.
\]

Hence, Theorem 2 tells us that for all \((\alpha_1, \beta_1)\) and \((\alpha_2, \beta_2)\), \( d(\cdot, \alpha_1, \beta_1) \) is more
dynamically inconsistent than \( d(\cdot, \alpha_2, \beta_2) \) if and only if \( \beta_1 < \beta_2 \). Note that the
values of the \( \alpha_n \) are irrelevant for the ranking of dynamic inconsistency.

As discussed towards the end of Section 4 and again towards the end of Section 6, for the utility weight function \( \lambda : A \to R_+^+ \), for which the solution to the welfare
maximization problem (2) is the equilibrium allocation, the function \( \lambda^{1/\gamma}/\rho : A \to R_+^+ \) may be a better approximation of the wealth share function \( \theta : A \to R_+^+ \) than
\( \lambda^{1/\gamma} : A \to R_+^+ \). It is thus important to identify the ranking of the measure
dynamic consistency when the weighted distributions \( \mu^2 \), defined from \( \lambda^{1/\gamma}/\rho \)
via (19), are Gamma distributions. To do so, suppose that the density function of
\( \mu^2 \) with respect to the Lebesgue measure coincides with (22). Then, by (20), the
density function of the corresponding \( \mu^1 \) with respect to the Lebesgue measure
Thus, the corresponding $\mu^1$ coincides with the Gamma distribution with parameters $(\alpha + 1, \beta)$. By the result of the previous paragraph, the discount factor function of which the distribution $\mu^2$ is the Gamma distribution $(\alpha_1, \beta_1)$ is more dynamically inconsistent that the discount factor function of which the distribution $\mu^2$ is the Gamma distribution $(\alpha_2, \beta_2)$ if and only if $\beta_1 < \beta_2$. Within the family of Gamma distributions, therefore, the ranking of the measures of dynamic inconsistency does not depend on whether the Gamma distributions are assumed for $\mu^1$ or $\mu^2$.

In concluding this subsection, we note that the Gamma distributions constitute an exponential family, and the the representative consumer’s discount rate functions are particularly easy to calculate for an exponential family. To see this, write $\zeta = -\beta$ and $\xi = \alpha - 1$, and define $w(\xi, q) = \xi \ln q$ and $v(\zeta, \xi) = (\xi + 1) \ln(-\zeta) - \Gamma(\xi + 1)$. Then

$$\frac{\beta^\alpha q^{\alpha - 1}}{\Gamma(\alpha)} \exp(-\beta q) = \exp(\zeta q + w(\xi, q) + v(\zeta, \xi)).$$

Thus, the Gamma distributions, indeed, constitute an exponential family. With a slight abuse of notation, denote by $K^1(\cdot, \zeta, \xi)$ the cumulant-generating function that corresponds to parameters $(\zeta, \xi)$ in the above expression, then $K^1(s, \zeta, \xi) = v(\zeta, \xi) - v(\zeta + s, \xi)$. Denote the corresponding discount rate function by $r(\cdot, \zeta, \xi)$. Then, by (9),

$$r(t, \zeta, \xi) = -\frac{\partial v}{\partial \zeta} \left( \xi - \frac{t}{\gamma}, \xi \right).$$

Thus, for an exponential family, Theorems 2 and 3 can be stated in terms of $v$.

---

3In fact, a family of distributions, parameterized by $(\zeta, \xi)$, that can be represented by the right-hand side of (24) is more general than an exponential distribution, because, in (24), $w(\xi, q)$
7.2 Bernoulli distributions

In this subsection, we consider the set of all Bernoulli distributions that put probability \(1/2\) to each of \(\alpha + \beta\) and \(\alpha - \beta\), where \(\alpha > \beta > 0\). We define the discount factor function \(d(\cdot, \alpha, \beta)\) and the discount rate function \(r(\cdot, \alpha, \beta)\) in the same way as we did in the previous subsection but using the Bernoulli distributions.

The cumulant-generating function \(K^1(\cdot, \alpha, \beta)\) is given by

\[
K^1(s, \alpha, \beta) = \ln \left( \frac{1}{2} \exp((\alpha + \beta)s) + \frac{1}{2} \exp((\alpha - \beta)s) \right).
\]

Hence,

\[
\frac{\partial K^1}{\partial s}(s, \alpha, \beta) = \alpha + \beta \frac{\exp(\beta s) - \exp(-\beta s)}{\exp(\beta s) + \exp(-\beta s)}, \tag{25}
\]

\[
\frac{\partial^2 K^1}{\partial s^2}(s, \alpha, \beta) = \frac{\beta^2}{(\exp(\beta s) + \exp(-\beta s))^2} \times \left( (\exp(\beta s) + \exp(-\beta s))^2 - (\exp(\beta s) - \exp(-\beta s))^2 \right)
\]

\[
= \left( \frac{2\beta}{\exp(\beta s) + \exp(-\beta s)} \right)^2. \tag{26}
\]

Since

\[
d \left( \frac{\beta}{\exp(\beta s) + \exp(-\beta s)} \right) = (\exp(\beta s) + \exp(-\beta s))^{-1} + \beta^2 (\exp(\beta s) + \exp(-\beta s))^{-2} (\exp(-\beta s) - \exp(\beta s)) > 0
\]

for every \(s \leq 0\), (26) is a strictly increasing function of \(\beta\). Thus the single-crossing property between the \(K^1(\cdot, \alpha_1, \beta_1)\) and \(K^1(\cdot, \alpha_2, \beta_3)\) and, hence, the single-crossing property between \(r(\cdot, \alpha_1, \beta_1)\) and \(r(\cdot, \alpha_2, \beta_2)\) stipulated in Theorem 3 hold if \(\beta_1 > \beta_2\).

As for the more-dynamically-inconsistent-than relation, note from (25) and (26) that \(\partial K^1(s, \alpha, \beta)/\partial s\) is a strictly increasing function of \(\alpha\) but \(\partial^2 K^1(s, \alpha, \beta)/\partial s^2\) does not depend on \(\alpha\). Hence, if \(\alpha_1 < \alpha_2\) and \(\beta_1 = \beta_2\), \(d(\cdot, \alpha_1, \beta_1)\) is more dynamically inconsistent than \(d(\cdot, \alpha_2, \beta_2)\). We will show that even if \(\alpha_1 = \alpha_2\) and \(\beta_1 > \beta_2\), need not be multiplicatively separable between \(\xi\) and \(q\).

\(^4\)At any \(s < 0\), the second term on the right-hand side of (25) is a strictly decreasing function of \(\beta\). Thus, if \((\alpha_1 - \alpha_2)(\beta_1 - \beta_2) \leq 0\), then \(\partial K^1(\cdot, \alpha_1, \beta_1)/\partial s\) and \(\partial K^1(\cdot, \alpha_2, \beta_2)/\partial s\) never intersect. Hence, if they do in fact intersect and satisfy the single-crossing property of Theorem 3, then \(\alpha_1 > \alpha_2\) and \(\beta_1 > \beta_2\), that is, the converse of the above claim holds.
\(d(\cdot, \alpha_1, \beta_1)\) is not more dynamically inconsistent than \(d(\cdot, \alpha_2, \beta_2)\). Note that in this case, the Bernoulli distribution with \((\alpha_2, \beta_2)\) second-order stochastically dominates the Bernoulli distribution \((\alpha_1, \beta_1)\). This fact shows, thus, that the second-order stochastic dominance relation is not sufficient for the more-dynamically-inconsistent-than relation.

By (25) and (26),

\[
\frac{\partial^2 K^1}{\partial s^2}(s, \alpha, \beta) = 4 \beta (\alpha + \beta) \exp(2\beta s) + (\alpha - \beta) \exp(-2\beta s) + 2\alpha)^{-2}
\]

Differentiate the curvature (27) with respect to \(\beta\), then we obtain

\[
((\alpha + \beta) \exp(2\beta s) + (\alpha - \beta) \exp(-2\beta s) + 2\alpha)^{-2}
\]

We claim that (28) is strictly positive for every \(s \leq 0\) sufficiently close to 0 but strictly negative for every sufficiently negative \(s\). First, if \(s = 0\), then (28) is equal to \(2\beta/\alpha\), which is strictly positive. Thus, (28) is strictly positive for every \(s \leq 0\) sufficiently close to 0. As for a sufficiently negative \(s\), since

\[
4\beta ((\alpha + \beta) \exp(2\beta s) + (\alpha - \beta) \exp(-2\beta s) + 2\alpha)^{-2}
\]

for every \(s\), it suffices to show that

\[
((2\alpha + \beta) - 2(\alpha + \beta)\beta s) \exp(2\beta s) + ((2\alpha - \beta) + 2(\alpha - \beta)\beta s) \exp(-2\beta s) + 4\alpha < 0 \quad (29)
\]

for every sufficiently large \(s\). To do so, note that the first term of the left-hand side of (29), \((2\alpha + \beta) - 2(\alpha + \beta)\beta s) \exp(2\beta s)\), converges to zero as \(s \to -\infty\). The second term, \((2\alpha - \beta) + 2(\alpha - \beta)\beta s) \exp(-2\beta s)\), diverges to \(-\infty\) as \(s \to -\infty\). The third term, \(4\alpha\), does not depend on \(s\). Thus, the left-hand side of (29) diverges.
to $-\infty$ as $s \to -\infty$, and, therefore, is strictly negative for every sufficiently negative $s$. This result implies that for all $\alpha$ and $\beta$ with $\alpha > \beta > 0$, a small increase in $\beta$ increases the curvature (27) for every $s \leq 0$ sufficiently close to 0 and decreases it for every sufficiently negative $s$. Thus, an increase in the standard deviation $\beta$ of the Bernoulli distribution, with its mean $\alpha$ fixed, does not make the cumulant-generating function more convex or more concave, and, by Theorem 2, does not make the representative consumer more dynamically inconsistent or less dynamically consistent.

A numerical example may be of some help to understand this somewhat inconclusive result. Take the mean 4% ($\alpha = 0.04$) and the standard deviation 3% ($\beta = 0.03$). Then the derivative (28) of the curvature (27) of the cumulant-generating function with respect to the standard deviation $\beta$ turns out to be positive if and only if $s$ is approximately less than 85. According to (18), this implies that if the coefficient of constant relative risk aversion is equal to one (the case of the logarithmic utility function) for both consumers, then a small (infinitesimal) increase in the standard deviation of the binary distribution of individual consumers' discount rates increases the local measure $-r'(t)/r(t)$ of the representative consumer’s dynamic inconsistency during the first 85 years, but such an increase decreases $-r'(t)/r(t)$ thereafter.\(^{6}\)

In concluding this section, we touch a generalization of the above analysis on the single-crossing property. Note that if $\varphi$ is a random variable taking values 1 and $-1$ with probability 1/2 each, then the distribution of its affine transformation $\alpha + \beta \varphi$ coincides with the Bernoulli distribution with parameter $(\alpha, \beta)$. Conversely, all Bernoulli distributions are generated by some affine transformations of $\varphi$. If, more generally, $\varphi$ is any random variable of zero mean that is bounded from below, then a family of the distributions of random variables $\alpha + \beta \varphi$ can be considered as a family of distributions of discount rates. An example of such families is the family of uniform distributions of which the supports are in $\mathbb{R}_+$.\(^{7}\) For such a family, $\partial K^1(s, \alpha, \beta)/\partial s$ is linear in $\alpha$, just like the right-hand side of (25). Hence, $d(\cdot, \alpha_1, \beta_1)$ is more dynamically inconsistent than $d(\cdot, \alpha_2, \beta_2)$ whenever $\alpha_1 < \alpha_2$ and $\beta_1 = \beta_2$.\(^{5}\)

\(^{5}\)The mean and standard deviation of this numerical example are chosen to match those used by Weitzman (2001, Sections III and IV).

\(^{6}\)In the context of climate change, 85 years is well within the range of decision making for which a careful choice of discount rates is required.

\(^{7}\)Take $\varphi$ so that it follows the uniform distribution, say, on $(-1,1)$. 

22
8 Term structure of interest rates

In this section, we explore the implication of Theorem 3 on the term structure of interest rates. In particular, we compare the yield curves, forward rates, and short-rate processes of the two economies of which the cumulant-generating functions of the approximate wealth-weighted distributions \(\mu^1\) of individual consumers’ discount rates have the single-crossing property.

8.1 Term structure of an economy

The state-price density process \(\pi\) is equal to the marginal utility process \(du'(e) = (d(t)u'(e(t)))_{t \in \mathbb{R}^+}\) evaluated at the aggregate consumption process \(e = (e(t))_{t \in \mathbb{R}^+}\). Hence the price at time \(t_1\), relative to the current consumption, of the discount bond with maturity \(t_2 > t_1\) is equal to

\[
E_{t_1} \left( \frac{\pi(t_2)}{\pi(t_1)} \right) = \frac{d(t_2)}{d(t_1)} E_{t_1} \left( \frac{u'(e_{t_2})}{u'(e_{t_1})} \right) = \exp \left( - \int_{t_1}^{t_2} r(t) \, dt \right) E_{t_1} \left( \frac{u'(e_{t_2})}{u'(e_{t_1})} \right). \tag{30}
\]

We denote this price by \(B(t_1, t_2)\). The yield to maturity, at time \(t_1\), of the discount bond with maturity \(t_2 > t_1\) is equal to

\[
-\frac{1}{t_1 - t_2} \ln B(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} r(t) \, dt - \frac{1}{t_2 - t_1} \ln E_{t_1} \left( \frac{u'(e_{t_2})}{u'(e_{t_1})} \right). \tag{31}
\]

We denote this \(g(t_1, t_2)\).

Another rate that we are interested in is the instantaneous forward rate, determined at time \(t_1\), for the delivery of the about-to-mature bond at time \(t_2\) is equal to

\[
-\frac{\partial}{\partial t_2} \ln B(t_1, t_2) = r(t_2) - \frac{d}{dt_2} \ln E_{t_1} \left( \frac{u'(e_{t_2})}{u'(e_{t_1})} \right),
\]

if

\[
E_{t_1} \left( \frac{u'(e_{t_2})}{u'(e_{t_1})} \right) \tag{32}
\]

is a differentiable function of \(t_2\). We denote this by \(f(t_1, t_2)\).

To define the short-rate process, we assume that the filtration \((\mathcal{F}(t))_{t \in \mathbb{R}^+}\) is generated by a one-dimensional Brownian motion \(B = (B(t))_{t \in \mathbb{R}^+}\) and the aggregate consumption process \(c\) is an \(\mathbb{R}_{++}\)-valued Itô process, written as

\[
de(t) = c(t)\mu(t) \, dt + c(t)\sigma(t) \, dB(t)
\]
for some $R$-valued adapted processes $\mu = (\mu(t))_{t \in \mathbb{R}_+}$ and $\sigma = (\sigma(t))_{t \in \mathbb{R}_+}$. Since $\pi = du'(c)$, Ito’s Lemma implies that $\pi$ is also an $\mathbb{R}_{++}$-valued Ito process, which can be written as

$$d\pi(t) = -\pi(t)h(t)\,dt - \pi(t)k(t)\,dB(t)$$

for some $R$-valued adapted processes $h = (h(t))_{t \in \mathbb{R}_+}$ and $k = (k(t))_{t \in \mathbb{R}_+}$. The process $h$ is the short-rate process, which gives an instantaneous riskless interest rate at each time $t$. Ito’s Lemma also implies that

$$h = r + \gamma \mu - \frac{\gamma(\gamma + 1)}{2} \sigma.$$

### 8.2 Comparison between two economies

Let $d_1$ and $d_2$ be the discount factor functions derived from two economies, of which the approximate wealth-weighted distributions are $\mu_1$ and $\mu_2$. That is, for each $n = 1, 2$,

$$d_n(t) = \left( \int_0^\infty \exp\left(-\frac{qt}{\gamma}\right) \, d\mu_n(q) \right)^\gamma.$$

Denote the cumulant-generating functions of $\mu_1$ and $\mu_2$ by $K_{11}$ and $K_{12}$. Let $r_1$ and $r_2$ be the corresponding discount rate functions. For each $n = 1, 2$, let $B_n$, $g_n$, $f_n$, and $h_n$ be the corresponding bond prices, yields to maturity, instantaneous forward rates, and short-rate process. Then

$$\frac{B_1(t_1, t_2)}{B_2(t_1, t_2)} = \exp\left(-\int_{t_1}^{t_2} (r_1(t) - r_2(t)) \, dt\right),$$

$$g_1(t_1, t_2) - g_2(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} (r_1(t) - r_2(t)) \, dt, \quad (34)$$

$$f_1(t_1, t_2) - f_2(t_1, t_2) = r_1(t_2) - r_2(t_2), \quad (35)$$

$$h_1(t) - h_2(t) = r_1(t) - r_2(t). \quad (36)$$

Note that while the $B_n(t_1, \cdot)$, $g_n(t_1, \cdot)$, $f_n(t_1, \cdot)$, and $r_n$ are, in general, stochastic processes (because, for each fixed $t_1$, (32) is a stochastic process in $t_2$), none of the above is stochastic. Moreover, the difference in the instantaneous forward rates, (35), does not depend on the time $t_1$ at which it is evaluated.

**Theorem 4** Suppose that for every $s \leq 0$, $(K_{11})''(s) > (K_{12})''(s)$ whenever $(K_{11})'(s) =
Thus, if whenever \( f_1(t_1, t_2) = f_2(t_1, t_2) \) whenever \( f_1(t_1, t_2) = f_2(t_1, t_2) \), then so does the instantaneous forward rates, that is, if the discount rates of two economies have the single-crossing property, then so does this theorem states that if the cumulant-generating functions of the distributions of discount rates of two economies have the single-crossing property, then so does the instantaneous forward rates, that is, if \( f_1(t_1, t_2) = f_2(t_1, t_2) \), then \( f_1(t_1, t) < f_2(t_1, t) \) for every \( t > t_2 \) and \( f_1(t_1, t) > f_2(t_1, t) \) for every \( t \in [t_1, t_2) \). In words, if the two economies have an equal instantaneous forward rate applicable to some future time, then the instantaneous forward rate applicable to any earlier time is higher, and that applicable to any later time is lower, in the first economy than in the second. Theorem 4 can be proved by using Theorem 3 and (35).

Theorem 4 compares the instantaneous forward rates, in the two economies, determined at a fixed time \( t_1 \) but with a variable delivery time \( t_2 \). Another comparison worth exploring is the instantaneous forward rates, with a fixed time to maturity, say \( \tau \), but with a variable time \( t_1 \) at which the rates are determined. What this means, in symbols, is how

\[
f_1(t_1, t_1 + \tau) - f_2(t_1, t_1 + \tau)
\]

depends on \( t_1 \), when \( \tau \) is a positive constant. In fact, by (35), (37) is equal to \( r_1(t_1 + \tau) - r_2(t_1 + \tau) \), which, when \( t_1 \) varies, behaves in the same way as \( r_1(t_2) - r_2(t_2) \) when \( t_2 \) varies. We can conclude, therefore, that the instantaneous forward rates in two economies, with a fixed time to maturity but with a variable time at which the rates are determined, also has the single-crossing property.

The next theorem shows the single-crossing property for the yields to maturity.

**Theorem 5** Suppose that for every \( s \leq 0 \), \((K_1^1)''(s) > (K_2^1)''(s) \) whenever \((K_1^1)'(s) = (K_2^1)'(s) \). Then, for every \( t_1 \geq 0 \) and every \( t_2 > t_1 \), \( \partial g_1(t_1, t_2)/\partial t_2 < \partial g_2(t_1, t_2)/\partial t_2 \) whenever \( g_1(t_1, t_2) = g_2(t_1, t_2) \).

**Proof of Theorem 5** By (34) and a straightforward calculation,

\[
\frac{\partial g_1}{\partial t_2}(t_1, t_2) - \frac{\partial g_1}{\partial t_2}(t_1, t_2) = \frac{1}{t_2 - t_1} \left((r_1(t_2) - r_2(t_2)) - (g_1(t_1, t_2) - g_2(t_1, t_2))\right).
\]

Thus, if \( g_1(t_1, t_2) = g_2(t_1, t_2) \), then

\[
\frac{\partial g_1}{\partial t_2}(t_1, t_2) - \frac{\partial g_1}{\partial t_2}(t_1, t_2) = \frac{1}{t_2 - t_1} (r_1(t_2) - r_2(t_2)),
\]

(38)
and, again by (34),
\[ \int_{t_1}^{t_2} (r_1(t) - r_2(t)) \, dt = 0. \]
Thus, there exists a \( t_0 \in (t_1, t_2) \) such that \( r_1(t_0) = r_2(t_0) \). By the single-crossing property of the \( K_n' \) and Theorem 2, the \( r_n \) also have the single-crossing property. Since \( t_2 > t_0 \), \( r_1(t_2) < r_2(t_2) \). By (38),
\[ \frac{\partial g_1}{\partial t_2}(t_1, t_2) - \frac{\partial g_1}{\partial t_2}(t_1, t_2) < 0. \]

The following theorem gives the relationship between the drift and diffusion terms of the two and the consequence of the single-crossing property of the derivatives of the cumulant-generating functions.

**Theorem 6** Suppose that for each \( n = 1, 2 \), the short-rate process \( h_n \) is an Ito process, written as
\[ dh_n(t) = \chi_n(t) \, dt + \eta_n(t) \, dB(t) \]
for some \( R \)-valued adapted processes \( \chi_n = (\chi_n(t))_{t \in R_+} \) and \( \eta_n = (\eta_n(t))_{t \in R_+} \). Then \( \chi_1 - \chi_2 = r_1' - r_2' \) and \( \eta_1 = \eta_2 \). Suppose, in addition, that for every \( s \leq 0 \), \( (K_1')''(s) > (K_2')''(s) \) whenever \( (K_1')'(s) = (K_2')'(s) \). Then for every \( t \geq 0 \), \( \chi_1(t) < \chi_2(t) \) whenever \( h_1(t) = h_2(t) \).

### 8.3 Heterogeneous CIR model

As an application of Theorem 6, we accommodate heterogeneous impatience into the model of the term structure of interest rates of Cox, Ingersoll, and Ross (1985). It is the only well known model of the term structure of interest rates that has a general equilibrium foundation, and, hence, allows us to investigate the impact of heterogeneous impatience on the term structure.

The original version of the CIR model has a representative consumer who exhibits constant relative risk aversion and constant discount rate. Thanks to the nature of the stochastic process defining the productivity of production technologies (Assumption 3 of Cox, Ingersoll, and Ross (1985)), the short-rate process exhibits mean reversion. To be specific, of the two economies we have been considering, we take the second economy as an economy of the CIR model with the
common coefficient $\gamma$ of constant relative risk aversion and the common discount rate $q$. Then there are a $\kappa^0 \in \mathbb{R}_{++}$, an $\bar{h} \in \mathbb{R}_{++}$, and a $\kappa^1 \in \mathbb{R}$ such that $\chi_2 = \kappa^0(\bar{h} - h_2)$ and $\eta_2 = \kappa^1\sqrt{h_2}$, that is,

$$dh_2(t) = \kappa^0(\bar{h} - h_2(t)) \, dt + \kappa^1\sqrt{h_2(t)} \, d\mathbb{B}(t).$$

(39)

Then, take the first economy as an economy in which all the individual consumers have the common coefficient $\gamma$ of constant relative risk aversion but do not have the same discount rate, and the aggregate endowment process coincides with that of the first economy. Then, by (17), $r_1' < 0 = r_2'$. Hence, by Theorem 6,

$$dh_1(t) = (\kappa^0(\bar{h} - h_2(t)) + r_1'(t)) \, dt + \kappa^1\sqrt{h_2(t)} \, d\mathbb{B}(t).$$

By (36),

$$dh_1(t) = \kappa^0 \left( \left( \bar{h} - q + r_1(t) - \frac{r_1'(t)}{\kappa^0} \right) - h_1(t) \right) \, dt$$

$$+ \kappa^1\sqrt{h_1(t) + q - r_1(t)} \, d\mathbb{B}(t).$$

(40)

This is a heterogeneous-impatience version of the CIR term structure of interest rate. Although the drift term of the short-rate process of the original version of the CIR model is reverting to the constant rate $\bar{h}$, the drift term of the heterogeneous-economy version is reverting to

$$\bar{h} - q + r_1(t) - \frac{r_1'(t)}{\kappa^0},$$

which is a function $t$. If, in addition, the approximate wealth-weighted distribution $\mu^1$ of individual consumers’ discount rates is the Gamma distribution with $\kappa^1 > 0$, then $\kappa^1 \gamma$. Since the CIR model is a model of a production economy, it cannot strictly be considered as a special case of the model of this paper. However, the equilibrium prices and consumption allocations of the CIR model coincide with those of the exchange economy of which the aggregate endowment process is nothing but the equilibrium aggregate consumption process of the CIR model. The preceding results are, therefore, applicable to the equilibrium prices of the CIR model.

8Since the original CIR model is a model of a production economy, incorporating heterogeneous impatience changes production decisions and, hence, the aggregate consumption process. The assumption of the common aggregate endowment process is, therefore, not strictly compatible with the equilibrium of the original CIR model. This assumption, however, simplifies the task of assessing the impact of heterogeneous impatience.
parameters \((\alpha, \beta)\), then, by (9) and (23),

\[
r_1(t) = \frac{\partial K^1}{\partial s} \left( -\frac{t}{\gamma}, \alpha, \beta \right) = \frac{\alpha \gamma}{t + \beta \gamma},
\]

\[
r'_1(t) = \frac{\partial^2 K^1}{\partial s^2} \left( -\frac{t}{\gamma}, \alpha, \beta \right) = -\frac{\alpha \gamma}{(t + \beta \gamma)^2}.
\]

Thus, the representative consumer in the first economy exhibits hyperbolic discounting. By (40),

\[
dh_1(t) = \kappa^0 \left( \bar{h} - q + \frac{\alpha \gamma}{t + \beta \gamma} - \frac{\alpha \gamma}{\kappa^0(t + \beta \gamma)^2} \right) - h_1(t) \, dt
\]

\[
+ \kappa^1 \sqrt{h_1(t) + q - \frac{\alpha \gamma}{t + \beta \gamma}} \, dB(t).
\]

## 9 Conclusion

In this paper, we have given a precise formulation to the notion that the more heterogeneous the individual consumers’ subjective discount rates are, the more dynamically inconsistent the representative consumer is. The measure of heterogeneity of discount rates is the convexity of the cumulant-generating function of the approximate wealth-weighted distribution of discount rates. We have also given two examples of parameterized families of distributions within which the measures of heterogeneity are compared, of which one consists of Gamma distributions and the other consists of Bernoulli distributions. We have applied these results to the analysis of the term structure of interest rates in heterogeneous economies. In particular, we have characterized the short-rate process in the version of the CIR in a heterogeneous economy.

There are a couple of points that have been left unanswered in this paper. First, we do not know the precise relationship between the utility weights in the maximization problem that defines the representative consumer’s discount rates and the wealth share at equilibrium. Second, we have not fully solved the equilibrium, at which the aggregate consumption levels are endogenously determined, of the heterogeneous-economy version of the CIR model. These two points should be clarified to assess the applicability of the results of this paper.

Most importantly, the analysis of this paper should be extended to the case where asset markets are incomplete. Since individual consumers have fewer in-
Instruments to transfer purchasing power across time and states, the impact of the heterogeneity of subjective discount rates on the representative consumer’s discount rates will be less pronounced than in the case of complete markets. To increase the relevance of the results of this paper, it is important to determine exactly how much the impact is reduced.

References


