Behavioral properties of the representative agent

Elyes Jouini* Clotilde Napp†

February 13, 2009

Abstract

In this paper, we show that behavioral features can be obtained at a group level when the
individuals of the group are heterogeneous enough. More precisely, starting from a standard
model of Pareto optimal allocations, with expected utility maximizers and exponential dis-
counting, but allowing for heterogeneity among agents’ beliefs and time preference rates, we
show that the representative agent exhibits interesting behavioral properties. In particular,
we obtain an inverse S-shaped probability distribution weighting function and hyperbolic
discounting. We provide possible interpretation as well as applications for this result.

JEL Codes : G11; D81; D84; D87; D03; H43

Keywords: behavioral agent, hyperbolic discounting, probability weighting function, rep-
resentative agent, neurofinance

1 Introduction

In this paper, we analyze efficient allocations in a model with von Neuman Morgenstern utility
maximizing agents and exponential discounting. Agents are heterogeneous, in the sense that
they might differ in their beliefs and in their time preference rates. At the aggregate level, the
social welfare function of this economy is characterized by a social/representative belief and a

*Université Paris Dauphine, Ceremade, F-75016 Paris. E-mail: jouini@ceremade.dauphine.fr
†CNRS, UMR 7088, F 75016 Paris
‡Université Paris Dauphine, DRM, F 75016 Paris
social/representative time preference rate. We examine these social characteristics and we show that we retrieve at the aggregate level, “behavioral” properties that have been proved to be true at the individual level in recent literature. The group acts as a "behavioral" agent and these behavioral properties at the aggregate level are generated by heterogeneity alone.

We first obtain that our heterogeneous agents can be aggregated into a representative agent. We find that the belief of the representative agent is essentially a mixture (or a power/Hölder average) of the individual beliefs\(^1\). In particular, since a mixture of Gaussian variables is not Gaussian, this implies that even if all agents anticipate normal distributions (with the same variance parameter but differ in the anticipated mean), the anticipated distribution at the aggregate level is not Gaussian. In the unbiased setting (i.e. in a setting where the average belief coincides with the objective belief), the mean remains the same but the variance is increased. More precisely, we show that the distribution of outcomes from the representative agent point of view is portfolio dominated by the objective distribution. This effect is reinforced when agents are more risk tolerant or when there is more heterogeneity.

As far as behavioral properties are concerned, we start by introducing natural notions of optimism and pessimism and we assume that beliefs are heterogeneous enough in order to allow for optimistic as well as pessimistic agents in the initial set of von Neuman Morgenstern utility maximizing agents. In such a setting, we obtain that the representative agent can neither be everywhere optimistic nor everywhere pessimistic; she is optimistic for the good states of the world and pessimistic for the bad states of the world. As in the SP/A Theory of Lopes (1987), the representative agent behaves as if she had fear (need for security) for very bad events and hope (desire for potential) for very good events. We obtain that the representative agent probability weighting function (i.e. the transformation of the objective decumulative distribution function into the decumulative distribution function of the representative agent) is inverse S-shaped as in Cumulative Prospect Theory. Moreover, we show that we are able to fit relatively well standard

\(^1\)This point has already been underlined by Shefrin (2005) and by Jouini-Napp (2007).
probability weighting functions of the Prospect Theory literature (Tversky and Kahneman, 1992, Tversky and Fox, 1995, Prelec, 1998, among others). We analyze how the distribution of individual characteristics in the group governs the shape of the resulting representative agent probability weighting function. More precisely, we show that attractiveness\(^2\) at the aggregate level is directly related to the average level of optimism while discriminability is related to beliefs heterogeneity.

As far as the consensus time preference rate is concerned, we obtain at the aggregate level a time preference rate that is lower than the average of the individual time preference rates and that is not constant over time. More precisely, the time preference rate of the representative agent appears as decreasing, which is consistent with “hyperbolic discounting”. It converges to the time preference rate of the most patient individual. These properties are similar to those obtained in a deterministic setting by Gollier-Zeckhauser (2005) and Lengwiler (2005) but are derived here in a stochastic setting that takes into account beliefs heterogeneity.

The implications of our results are twofold. On the one hand, we obtain that a ”behavioral” individual (i.e. an individual whose preferences are governed by hyperbolic discounting and Cumulative Prospect Theory) behaves as would, at a Pareto optimum, a group of standard heterogeneous vNM individuals with exponential discounting. Our results can be related to a recent strand of research, called Neuroeconomics. The purpose of neuroeconomic theory is to take a look into the brain (that is considered as a black box in economic theory including behavioral economic theory) and produce a novel theory of individual decision-making based on experimental neurosciences and/or on models of agents interactions. Developments in brain imaging have helped identify different regions of the brain that are associated to different types of processes, different time preferences, different information processing, etc. If each region or each process is represented by an agent, Economic Theory provides then many useful tools

\(^2\)Gonzalez and Wu (1999) propose a typology of probability weighting functions based on the concepts of attractiveness and discriminability.
to analyse the brain. As underlined by Brocas and Carillo (2008), "the (...) goal of this strand of research is to revisit the individual decision-making paradigm and provide micro-microfoundations for characteristics of human behavior that have been traditionally ignored or considered as exogeneously given". Examples include hyperbolic discounting, distorted beliefs, mental accounting, etc. For instance, Brocas and Carillo (2008) analyse interactions between different brain areas through principal-agent models.

Our model may be interpreted as an optimal resources allocation model between different brain areas. Our conclusions can then be rephrased as follows: a model of the brain with a central planner (the cortex) who relies on evaluations provided by doers (mental processes with heterogeneous time preference rates and beliefs) in order to evaluate risky prospects leads to hyperbolic discounting and to probability weighting functions that have the same shape as in the Cumulative Prospect Theory (CPT).

On the other hand, our results imply that in order to analyse the properties of Pareto optima with a group of individuals endowed with “behavioral” beliefs and hyperbolic discounting, one only needs to consider the properties of Pareto optima within a larger group of individuals endowed with heterogeneous standard beliefs and time preference rates³.

Note that we don’t pretend to retrieve all features of CPT on the aggregate belief nor all features of the time preference rate as in e.g. Loewenstein and Prelec (1992). We only retrieve one of the three main features of CPT, namely the inverse S-Shaped probability distribution weighting function (the other two being the presence of a reference point and the presence of loss aversion). This comes from the fact that we have introduced heterogeneity on the beliefs only, hence the ”behavioral” property that we retrieve deals with the belief only. We also only obtain the ”hyperbolic” property of the time preference rate and not other behavioral properties such as the different (discounting) treatment of gains and losses. In order to retrieve such properties

³Models with heterogeneous beliefs and time preference rates have been studied by, e.g., Jouini and Napp (2007).
it likely would be necessary to introduce heterogeneity on the utility functions.

The paper is organised as follows. Section 2 presents the model. Section 3 analyses the properties of the belief of the representative agent, while Section 4 analyses the properties of the time preference rate of the representative agent. Section 5 provides possible interpretations as well as applications and concludes.

All proofs are in the e-companion.

2 The Model

We consider an economy with a single consumption good and with agents that have the same utility function and heterogeneous beliefs. Aggregate endowment in the consumption good is described by a random variable $e^*$ defined on a probability space $(\Omega, F, P)$. We let $I$ denote the set of heterogeneous agents. We assume that the common utility function is given by $u(x) = \frac{x^{1-\frac{1}{\beta}}}{1-\frac{1}{\beta}}$. Each agent has a subjective belief $Q_i$ about the distribution of $e^*$ and wants to maximize her von Neumann Morgenstern utility for consumption of the form $U_i(c) = E^{Q_i}[u(c)]$. We let $M^i$ denote the density of $Q_i$ with respect to the objective probability $P$, hence agent $i$’s utility for consumption can equivalently be written in the form $U_i(c) = E[M^i u(c)]$.

In such an economy, we consider the aggregate utility function $U$ defined as the solution of the following maximization program

$$
U(e^*) = \max_{\sum_{i=1}^{\#i} y^i = e^*} \sum_{i=1}^{\#i} \lambda_i E[M^i u(y^i)]
$$

where $(\lambda_i)$ are given positive weights. The aggregate utility function corresponds to the value of the social welfare function at the Pareto optimum when agent $i$ is granted a weight $\lambda_i$ by a social planner. The index $i$ may also represent a group of agents with common beliefs $M^i$ and $y^i$ then represents the total consumption of the group and $\lambda_i$ the sum of the weights granted by the social planner to the individuals in the group. When the social planner grants the same
weight to all the agents in the economy, the weight \( \lambda_i \) represents the proportion of agents that have the same belief \( M^i \). From a social planner point of view, the aggregate utility function corresponds to the highest social utility level among all possible endowment distributions across agents.

The number of agents can be finite or infinite. In the case of an infinite number of agents, sums are replaced by integrals. We obtain the following representation result.

**Proposition 1 Representative Agent**

The aggregate utility for consumption is given by

\[
U(e^*) = E [M u(e^*)]
\]

with

\[
M = \left( \sum_{i \in I} \gamma_i (M^i)^{\frac{1}{\eta}} \right) \frac{1}{\eta}
\]  

(1)

for \( \gamma_i = \lambda_i^{\frac{1}{\eta}} \). The representative agent belief is then given by \( M = \left( \sum_{i \in I} \gamma_i (M^i)^{\frac{1}{\eta}} \right) \frac{1}{\eta} \).

This means that, at the Pareto optimum, the aggregate utility is given by the utility of a representative agent endowed with an average belief (and the same utility function as each of the agents). In particular, if all the agents share the same belief, then the representative agent will share this common belief. If we think of \( e^* \) as a given prospect for the group \( I \) of agents, the aggregate utility \( U(e^*) \) corresponds to the social welfare associated with the optimal allocation of \( e^* \) across the members of the group and is given by the utility of the representative agent. Consequently, if the group has to choose between two different prospects respectively denoted by \( e^1 \) and \( e^2 \), the prospect \( e^1 \) will be preferred to the prospect \( e^2 \) if and only if the representative agent with belief \( M \) prefers \( e^1 \) to \( e^2 \). The optimal decisions of the group are given by the decisions of the representative agent with belief \( M \).

In the next section we analyse the properties of the representative agent belief, since it is a
key determinant of the decisions of the group.

3 Aggregate endowment distribution from the representative agent point of view

We start by analysing the properties of the distribution of aggregate endowment from the representative agent point of view. In particular, we want to determine its expression and how it relates to the individual subjective distributions. If the individual distributions are lognormal, is the representative agent’s distribution also lognormal? How are the mean and the variance of the representative agent’s distribution, compared to the mean and variance of the individual distributions? Can we exhibit stochastic dominance properties?

Let us first introduce some technical definitions and notations. As underlined by Jouini-Napp (2007) among others, we don’t have $E[M] = 1$ and $M$ fails to be the density of a probability measure (except in the specific logarithmic utility setting). We say that the distribution of a random variable $X \equiv \varphi(e^*)$ admits a “density $f_X$ for the representative agent” if for all function $h$, we have $E[Mh(X)] = \int h(x)f_X(x)\,dx$. Moreover, in order to analyse the relative weights of the different states of the world from the representative agent point of view, we introduce the normalized belief $\overline{M} \equiv \frac{M}{E[M]}$ and its associated probability measure $Q$ defined by $\frac{dQ}{dP} \equiv \overline{M}$.

We suppose that for all $i \in I$, the distribution of $e^*$ for agent $i$ admits a density$^4$ (with respect to the Lebesgue measure on the real line), denoted by $f^i$. We also assume that the objective distribution of $e^*$ (i.e. the distribution of $e^*$ under the objective probability $P$) admits a density and we denote it by $f$.

**Proposition 2** The distribution of $e^*$ admits the following density for the representative agent

$$f^M = \left( \sum_{i \in I} \gamma_i \left( f^i \right)^\eta \right)^{1/\eta}$$

$^4$In other words, the distribution of $e^*$ under $Q_i$ is absolutely continuous with respect to the Lebesgue measure.
which is a power average of the initial densities. In particular, for $\eta = 1$, the distribution of $e^*$ for the representative agent is a mixture of the individual subjective distributions.

We have already seen through Proposition 1 that the representative agent belief $M$ is an average of the individual beliefs $M^i$, analogously Proposition 2 states that the density $f^M$ of $e^*$ for the representative agent is an average of the densities $f^i$ of $e^*$ for the different agents. As an immediate consequence of Proposition 2, we get that for any measurable real-valued function $\varphi$, the distribution of $\varphi(e^*)$ admits the density $f^{M, \varphi} = \left(\sum_{i \in I} \gamma_i \left(f^{i, \varphi}\right)^\eta\right)^{1/\eta}$ for the representative agent where $f^{i, \varphi}$ denotes the density of the distribution of $\varphi(e^*)$ for agent $i$. This implies in particular that in the case $\eta = 1$, if each agent anticipates a normal distribution on $\log e^*$, then the distribution of $\log e^*$ is a mixture\footnote{Kon (1984) offers a pioneering work in the introduction of mixture models with the use of the mixture of two normal distributions as an alternative characterization of US stock returns. Subsequent studies suggest the adequacy of such mixtures for different markets (see Söderlind, 1997, for interest rate markets, Booth and Glassman, 1987, for exchange rate models, and Pan, Chan and Fok, 1995 for future markets).} of normal distributions.

Consider the implications in terms of mean and variance, in the setting with $\eta = 1$. We have

$$E^Q[e^*] = \sum_{i \in I} \gamma_i E^{Q_i}[e^*].$$

This means that the mean at the aggregate level is given by an arithmetic average of the individual means.

As far as the variance is concerned, we have

$$Var^Q[e^*] = \sum_{i \in I} \gamma_i Var^{Q_i}[e^*] + Var_i \left(E^{Q_i}[e^*]\right),$$

where $Var_i \left(E^{Q_i}[e^*]\right)$ measures beliefs (on the mean) heterogeneity and is given by $Var_i \left(E^{Q_i}[e^*]\right) = \sum_{i \in I} \gamma_i \left(E^{Q_i}[e^*]\right)^2 - \left(\sum_{i \in I} \gamma_i E^{Q_i}[e^*]\right)^2$. This means that the variance at the aggregate level is given not only by an arithmetic average of the individual variances, but also by an additional term related to beliefs dispersion. The variance is “increased” at the aggregate level and this
increase is proportional to the level of beliefs heterogeneity.

In particular, even if the agents agree on the objective variance, i.e. when \( \text{Var}^{Q_i}[e^*] = \text{Var}^P[e^*] \), and if the average of the individual means is equal to the objective mean, we have \( E^Q[e^*] = E^P[e^*] \) and \( \text{Var}^Q[e^*] = \text{Var}^P[e^*] + \text{Var}_i \left( E^Q_i[e^*] \right) \), which means that the mean is unchanged but that there is more variance at the aggregate level than at the objective level: beliefs heterogeneity generates “doubt”.

Let us consider more precisely the case of lognormal distributions. We assume that the distribution of aggregate endowment is lognormal, with \( \log e^* \sim \mathcal{N}(\mu, \sigma^2) \) under the objective probability \( P \). Agents agree about the variance but disagree about the mean of the random variable \( e^* \), i.e., we assume that under \( Q_i \) the distribution of \( \log e^* \) is given by \( \mathcal{N}(\mu_i, \sigma^2) \) with density \( f^{\log}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \). When \( \mu_i \) is smaller than \( \mu \), we will say that agent \( i \) is pessimistic and when \( \mu_i \) is larger than \( \mu \), we will say that agent \( i \) is optimistic.

**Proposition 3** Consider the setting with two equally weighted groups of agents and lognormal subjective distributions \( \ln \mathcal{N}(\mu_i, \sigma^2) \), for \( i = 1, 2 \).

1. The distribution of \( \log e^* \) for the representative agent is not Gaussian. When \( \Delta \equiv |\mu_1 - \mu_2| > \frac{2\sigma}{\sqrt{\eta}} \), the distribution of \( \log e^* \) is bimodal.

2. For \( \eta = 1 \), the distribution of \( \log e^* \) for the representative agent has the following moments

\[
E^Q[\log e^*] = \frac{\mu_1 + \mu_2}{2},
\]

\[
\text{Var}^Q[\log e^*] = \text{Var}^P[\log e^*] + \text{Var}_i(\mu_i) = \sigma^2 + \frac{\Delta^2}{4},
\]

\[
\beta^2_2 = 3 - \frac{2\Delta^4}{(\Delta^2 + 4\sigma^2)^2} < 3,
\]

where \( \beta^2_2 \equiv \frac{E^Q[(\log e^* - E^Q(\log e^*))^4]}{E^Q[(\log e^* - E^Q(\log e^*))^2]^2} \) is the kurtosis excess coefficient.

3. For general \( \eta \) and when \( \mu = \frac{\mu_1 + \mu_2}{2} \), we have \( \frac{dQ}{dP} = h(e^*) \) where \( h \) is symmetric with respect to \( \mu \), decreasing before \( \mu \) and increasing after \( \mu \). In particular, we have \( E^Q[\log e^*] = \)
$E^P[\log e^*], \ Var^Q[\log e^*] > Var^P[\log e^*]$ and $Q$ puts more weight on the tails of the distribution of $\log e^*$ than $P$.

4. For general $\eta$ and general $(\mu_i)$, we have $E^Q[\log e^*] = \frac{\mu_1 + \mu_2}{2}$ and $Var^Q[\log e^*] > Var^P[\log e^*]$.

5. For $\eta > \eta'$ and associated representative agent probability measures $Q^\eta$ and $Q'^\eta$, we have

$$\frac{dQ^\eta}{dQ'^\eta} = h_{\eta, \eta'}(e^*)$$

where $h_{\eta, \eta'}$ is symmetric with respect to $\frac{\mu_1 + \mu_2}{2}$, decreasing before $\frac{\mu_1 + \mu_2}{2}$, and increasing after $\frac{\mu_1 + \mu_2}{2}$. In particular, $Var^Q[\log e^*]$ increases with the level or risk tolerance $\eta$.

The first point of the Proposition means that the distribution of $\log e^*$ is not Gaussian even if the individual distributions are Gaussian. For $\eta = 1$, this reduces to the well known fact that a mixture of normal distributions is not normal. The bimodality of the distribution of $\log e^*$ is illustrated on Figure 1.

Points 2, 3, and 4 imply that when there are optimistic as well as pessimistic agents in the group, the distribution of $\log e^*$ for the representative agent has more variance and is more flat (platikurtic) than the objective distribution. It also has more variance than each of the individual subjective distributions. Roughly speaking, these results imply that difference in beliefs induces more risk at the representative agent level. When there is no bias, i.e., when $\frac{\mu_1 + \mu_2}{2} = \mu$, the mean is unchanged. Figure 1 and 2 illustrate these conclusions in different settings. Note that Figure 1 is similar to Figure 8.2 in Shefrin (2005).

Point 3 is more than a result on the variance. It implies that the representative agent’s distribution puts more weight on the tails. More precisely, it implies that the distribution of $\log e^*$ for the representative agent is Portfolio Dominated by the objective distribution. Let us recall that a distribution $f$ dominates a distribution $g$ in the sense of Portfolio Dominance ($f \succeq_{PD} g$) if we have $\int u'(x)(x - a)f(x)dx = 0 \implies \int u'(x)(x - a)g(x)dx = 0$ for any real number $a$ and any non-decreasing concave function $u$. This concept has been introduced in the context of portfolio problems by Landsberger and Meilijson (1993) and further studied by
Gollier (1997). It characterizes the changes in the distribution of the returns of the risky asset that lead to an increase in demand for the risky asset irrespective of the risk-free rate. It is then related to the degree of riskiness. Hence, aggregate endowment is considered as more risky by the representative agent than it actually is. The variance properties are easily deduced from the portfolio dominance properties, since as shown in Jouini and Napp (2008), a mean preserving PD dominated shift for a given distribution increases the variance.

Point 5 of the proposition shows that, for given individual distributions, a higher level of risk tolerance induces a portfolio dominated shift in the representative agent’s distribution. The interpretation is as follows. When there is heterogeneity, the individual optimal allocations exhibit more variance than in the homogeneous setting. Indeed, each agent will consume a larger proportion of aggregate endowment in states of the world that she considers more likely, and a smaller proportion of aggregate endowment in states of the world that the other agent considers more likely. However, this effect induced by beliefs heterogeneity is counterbalanced by risk aversion. Consequently, the higher the level of risk tolerance, the more heterogeneous the members of the group are in their optimal allocations. Risk aversion and heterogeneity impact individual allocations, and aggregate belief, in opposite directions. Figure 2 illustrates this result. In particular, one can see on Figure 2 that an increase in the level of risk tolerance increases the distance between the peaks and their size. The maximal distance corresponds to an infinite level of risk tolerance; the first (resp. second) peak of the representative agent belief corresponds to the peak of the first (resp. the second) agent’s subjective belief.

4 Behavioral properties of the representative agent

The discussion following Proposition 2 as well as Proposition 3 suggest that the representative agent exhibits interesting behavioral properties, similar to those introduced in the recent behavioral economics literature (Cumulative Prospect Theory of Kahneman and Tversky, 1992, or SP/A Theory of Lopes, 1987). This section analyses these similarities more precisely.
We start by introducing the notion of pessimism/optimism. For lognormal distributions, there is a natural order on the set of possible distribution functions induced by the natural order on the means \( \mu_i \). As we have underlined it, agents with a larger (resp. smaller) \( \mu_i \) can be referred to as more optimistic (resp. pessimistic). In a more general setting, we introduce the following notion of optimism/pessimism.

**Definition 1** An agent is said to be (everywhere) optimistic (resp. pessimistic) if \( \frac{f_i}{f_j} \) is non-decreasing (resp. nonincreasing). Agent \( i \) is said to be more optimistic than agent \( j \) and we denote it by \( f_i \preceq f_j \) if and only if \( \frac{f_i}{f_j} \) is nondecreasing. The relation \( \preceq \) is an order on the set \( (f_i)_{i \in I} \).

Definition 1 can be rephrased in terms of Monotone Likelihood Ratio Dominance (MLR): agent \( i \) is more optimistic than agent \( j \) if the distribution of \( e^* \) for agent \( i \) (i.e., under \( Q_i \)) dominates the distribution of \( e^* \) for agent \( j \) (i.e., under \( Q_j \)) in the sense of the MLR. For a given agent \( i \), if we let \( g_i \) denote the transformation of the objective decumulative distribution function \( F \) into the agent’s subjective decumulative distribution function \( F_i \), i.e. such that \( F_i = g_i \circ F \), it is easy to check that \( \frac{F_i}{F_j} \) is nondecreasing (resp. nonincreasing) if and only if \( g_i \) is convex (resp. concave). This means that our concept of optimism/pessimism is the analog, in the expected utility framework, of the concept of optimism/pessimism introduced by Diecidue and Wakker (2001) in a RDEU framework. Other concepts of optimism/pessimism have been proposed in the literature. In particular, Yaari (1987), Chateauneuf and Cohen (1994) and Abel (2002) propose a definition based on First Stochastic Dominance. Note that MLR dominance is stronger than FSD.

A MLR dominated shift for a given distribution reduces the mean and if agent \( i \) is more

---

6This concept is widely used in the statistical literature and was first introduced in the context of portfolio problems by Landsberger and Meilijson (1990). More precisely, Landsberger and Meilijson (1990) showed that in the standard portfolio problem a MLR shift in the distribution of returns of the risky asset leads to an increase in demand for the risky asset for all agents with nondecreasing utilities.

7More precisely, in an expected utility framework Abel (2002) defines pessimism by the condition \( F_i \geq F \) (First Stochastic Dominance) that corresponds to the condition \( g_i \geq lD \) introduced by Chateauneuf and Cohen (1994) in a RDEU setting.
pessimistic than agent \( j \) we have \( E_{Q_i} [e^*] \leq E_{Q_j} [e^*] \). This last condition characterizes the MLR dominance when we restrict our attention to a family of lognormal distributions with the same variance parameter and we then retrieve that agent \( i \) is more optimistic than agent \( j \) if and only if \( \mu_i > \mu_j \). In that case, optimistic (resp. pessimistic) agents are then characterized by \( \mu_i > \mu \) (resp. \( \mu_i < \mu \)) as in Shefrin (2005).

**Proposition 4** Behavioral properties of the representative agent belief.

We assume that there are at least one optimistic agent denoted by \( f_{\text{opt}} \) and one pessimistic agent denoted by \( f_{\text{pess}} \). We also assume that \( \lim_{+\infty} f_{\text{opt}} = \lim_{-\infty} f_{\text{pess}} = +\infty \) and \( \lim_{-\infty} f_{\text{opt}} = \lim_{+\infty} f_{\text{pess}} = 0 \).

1. The representative agent can neither be (everywhere) optimistic, nor (everywhere) pessimistic, i.e. \( f^M \) is non monotone.

2. The representative agent overestimates the weight of the “good states of the world” (high values of \( e^* \)) as well as the weight of the “bad states of the world” (low values of \( e^* \)), i.e. there exist \( x_{\text{inf}} \) and \( x_{\sup} \) such that \( f^M(x) \geq 1 \) for \( x \leq x_{\text{inf}} \) and \( f^M(x) \geq 1 \) for \( x \geq x_{\sup} \).

3. The representative agent acts as if she had fear (need for security) for very bad events and hope (desire for potential) for very good events, i.e. there exist \( x_{\text{inf}} \) and \( x_{\sup} \) such that \( Q(x \leq \underline{x}) \geq P(x \leq \underline{x}) \) for \( \underline{x} \leq x_{\text{inf}} \) and \( Q(x \geq \overline{x}) \geq P(x \geq \overline{x}) \) for \( \overline{x} \geq x_{\sup} \).

4. If one of the agents denoted by \( f_{\text{opt}}^{\max} \) is more optimistic than all the other agents and if one of the agents denoted by \( f_{\text{pess}}^{\max} \) is more pessimistic than all the other agents, then the representative agent behaves like the most pessimistic individual for low values of \( e^* \) and behaves like the most optimistic investor for high values of \( e^* \), i.e. \( f^M \sim_{-\infty} f_{\text{opt}}^{\max} \) and \( f^M \sim_{+\infty} f_{\text{pess}}^{\max} \).

By definition, \( f_{\text{opt}} \) (resp. \( f_{\text{pess}} \)) is nondecreasing (resp. nonincreasing). In Proposition 4, we slightly reinforce these conditions by further assuming that the values of \( f_{\text{opt}} \) (resp. \( f_{\text{pess}} \))
range from zero to infinity. Notice that these conditions are satisfied in the case of lognormal distributions.

It appears from this proposition as long as there are optimistic as well as pessimistic agents, the representative agent behaves like the individual agents considered in the behavioral economics and/or psychology literature. Indeed, she puts more weight on small probability events with large consequences as in the Cumulative Prospect Theory of Kahneman and Tversky (1992), has fear (need for security) for very bad events and hope (desire for potential) for very good events as in the SP/A Theory of Lopes (1987). Everything works then as if the representative agent distorted the objective probability as in the Cumulative Prospect Theory. In the next section, we analyze more precisely how this distortion operates, through the probability weighting function.

5 Probability weighting function

Let us first recall that in the context of Cumulative Prospect Theory, the probability weighting function $w$ is defined as the function that transforms the decumulative objective distribution function into the decumulative subjective distribution function $^8$. We first analyse if the probability weighting function of the representative agent has the same shape as in the Cumulative Prospect Theory. We then relate the properties of the probability weighting function to the individual beliefs of the agents, and in particular how shifts in the individual characteristics impact the probability weighting function.

5.1 Inverse S-Shape of the probability weighting function

A variety of methods have been used to determine the shape of the probability weighting function. Tversky and Kahneman (1992), Fox and Tversky (1995) and Prelec (1998) among others specify parametric forms (respectively $\omega (p) = \frac{p^\gamma}{[p^\gamma+(1-p)^\gamma]^{1/\gamma}}, \omega (p) = \frac{\delta p^\gamma}{[\delta p^\gamma+(1-p)^\gamma]^{1/\gamma}}$ and

$^8$If $g$ denotes the objective density and if $h$ denotes the subjective density, we have $w(\int_{x}^{\infty} g(s)ds) = \int_{x}^{\infty} h(s)ds$. 

\[14\]
\[ \omega(p) = \exp(-\log p^\gamma) \] and estimate them through standard techniques. Wu and Gonzalez (1996, 1998) and Abdellaoui (2000) avoid the potential problems of parametric estimation and directly derive from experimental studies the shape of the probability weighting function at the aggregate or individual level.

The results of all these studies are (mostly) consistent with an inverse S-shaped weighting function, concave for small probabilities, and convex for moderate and high probabilities. As it is shown in the next proposition, these properties are retrieved at the representative agent level.

**Proposition 5**

1. In the lognormal setting, if the set \( I \) is made of both optimistic and pessimistic agents then the transformation of the objective decumulative distribution function into the decumulative distribution function of the representative agent is inverse S-shaped: concave then convex.

2. In the general setting, if there are at least one optimistic agent \( f_{opt} \) and one pessimistic agent \( f_{pess} \) with \( \lim_{-\infty} f_{opt} = \lim_{+\infty} f_{pess} = +\infty \) and \( \lim_{-\infty} f_{opt} = \lim_{+\infty} f_{pess} = 0 \) and if the probability weighting function is continuously twice differentiable on \([0, 1]\), then it is inverse S-shaped at the boundaries of \([0, 1]\): concave for small probabilities, and convex for high probabilities.

Figure 3 represents the representative agent probability weighting function in a model with two logarithmic utility agents and illustrates the results of Proposition 5.

Moreover, Figure 4 permits to show that with a well chosen distribution of agents’ characteristics we obtain a probability weighting function that perfectly fits Prelec (1998)’s function.

In the lognormal setting, if we denote by \( \delta_i \) the quantity \( \delta_i = \frac{\mu - \mu}{\sigma} \), it is interesting to remark that the probability weighting function of the representative agent only depends on the \( \delta_i \)s and on the relative proportions \( \gamma_i \)s and is independent of \( \mu \) and \( \sigma \). In other words, the probability weighting function of the representative agent only depends on how much the agents deviate
from the objective mean in terms of standard deviation.

5.2 Discriminability, attractiveness and individual agents’ beliefs

Let us analyse how the main features for the shape of the probability weighting function at the aggregate level are related to the individual characteristics of the agents. Gonzalez and Wu (1999) exhibit two main features for the shape of the probability weighting function: diminishing sensitivity and attractiveness.

Attractiveness characterizes the absolute level of the probability weighting function. Indeed, an inverse S-shaped function can be completely below the identity line, can cross the identity line at some point or can be completely above the identity line. If an agent has a probability weighting function graph that is more “elevated” than the probability weighting function graph of another agent, then this means that the first agent finds betting on the chance domain more attractive than the second agent.

More precisely, the definition of attractiveness can be expressed in terms of First Stochastic Dominance (FSD). If we denote by $g_i$ the probability weighting function of agent $i$ then agent $i$ finds betting on the chance domain more attractive than agent $j$ if $g_i \geq g_j$ or, in other words, if $F_i \leq F_j$ which characterizes the First Stochastic Dominance. We will then say that the probability weighting function $g_i$ is more attractive than the probability weighting function $g_j$ when $f_i$ dominates $f_j$ in the sense of the First Stochastic Dominance.

Let $(\gamma_i)$ and $(\gamma'_i)$ denote two possible distributions of agents’ characteristics; as usual, we will say that the distribution $(\gamma'_i)$ dominates the distribution $(\gamma_i)$ in the sense of the FSD if for any increasing family $(x_i)$, we have $\sum \gamma_i x_i \leq \sum \gamma'_i x_i$. In other words, the distribution $(\gamma'_i)$ puts more weight on more attractive distributions. We will say that the set $(f_i)_{i \in I}$ of agent’s density functions is totally ordered with respect to the MLR order, i.e. for all $(i, j)$ we have either $f_i \succeq_{MLR} f_j$ or $f_j \succeq_{MLR} f_i$ and we will say that the distribution $(\gamma'_i)$ dominates the distribution $(\gamma_i)$ in the sense of the MLR if whenever $f_i \succ_{MLR} f_j$ we have $\frac{\gamma'_i}{\gamma_i} \geq \frac{\gamma'_j}{\gamma_j}$. In other words the ratio
between the two densities \( (\gamma'_i) \) and \( (\gamma_i) \) increases with agents’ optimism and, in particular, the distribution \( (\gamma'_i) \) puts more weight on more optimistic agents.

**Proposition 6**

1. For log-utility functions and in the case of lognormal distributions with the same variance parameter \( \sigma^2 \), a FSD shift on the distribution of the means \( (\mu_i) \) of the agents’ density functions increases attractiveness of the representative agent’s probability weighting function.

2. For log-utility functions and general distributions, if the set \( (f_i)_{i \in I} \) of agent’s density functions is totally ordered with respect to the FSD order then a FSD shift in the distribution of agents’ density functions \( (f_i)_{i \in I} \) increases attractiveness of the representative agent’s probability weighting function.

3. For general CARA utility functions and general distributions, if the set \( (f_i)_{i \in I} \) of agent’s density functions is totally ordered with respect to the MLR order then a MLR dominated shift in agents’ density functions distribution leads to a more pessimistic representative agent.

When all agents have logarithmic utility functions, attractiveness at the representative agent level increases when the weight granted to the more attractive density functions increases. Since FSD is weaker than MLR, attractiveness at the representative agent level increases when the weight granted to the more optimistic agents increases. This is illustrated by Figure 5. As shown in the previous proposition, this last property can be extended to power utility functions if we replace FSD shifts on the distribution of agents’ characteristics by MLR shifts.

Diminishing sensitivity corresponds to the fact that people become less sensitive to changes in probability as they move away from a reference point. In the probability domain, the two endpoints 0 (certainly will not happen) and 1 (certainly will happen) serve as reference points and under this principle, increments near the endpoints of probability loom larger than increments near the middle of the scale. This concept is related to the concept of discriminability.
in psychophysics literature and can be illustrated by two extreme cases: a function that approaches a step function and a function that is almost linear. In the next proposition we show that the level of discriminability at the representative agent level is closely related to the level of disagreement among agents.

Let us consider as above a family of agents with lognormal distributions $\ln \mathcal{N}(\mu_i, \sigma^2)$. We denote by $(\mu_i)$ the support of the distribution of the mean parameter and by $(\gamma_i)$ the associated weights. Recall that a mean preserving spread is defined as a modification of the distribution of $(\gamma_i)$ on a set of three locations $\mu_1 < \mu_2 < \mu_3$ with associated increments $\delta_1 \geq 0, \delta_2 \leq 0$ and $\delta_3 \geq 0$ such that $\sum_{i=1}^{3} \delta_i = 0$ and $\sum_{i=1}^{3} \delta_i \mu_i = 0$. A mean preserving spread will be said symmetric if $\delta_1 = \delta_3$.

**Proposition 7** For log-utility functions and in the case of lognormal distributions $\ln \mathcal{N}(\mu_i, \sigma^2)$, a symmetric mean-preserving spread on the distribution of $(\mu_i)$ decreases discriminability.

Intuitively, this proposition means that when the level of disagreement among agents increases, then the representative agent focuses more on the endpoints of the probability domain and is less sensitive to probability variations in the middle of the scale. Figure 6 illustrates this result. It shows, in the setting with two agents, that discriminability decreases with the level of disagreement. When both agents agree on the objective distribution, the probability weighting function is linear. When the agents disagree, one of them overestimating the average payoff by twice the standard deviation and the other one underestimating it by twice the standard deviation, we obtain a function that approaches a step function.

6 The Setting with Heterogeneous Time Preference Rates

In this section, we extend our framework in order to take into account the impact of time and heterogeneous time preference rates across the agents. Aggregate endowment at a given date $t$ is described by a random variable $e_t^*$. Agents have different time preference rates $(\rho_i)$ and...
different subjective beliefs $Q_i$ about the distribution of $e_i^t$. We let $M_i^t$ denote the density at date $t$ of $Q_i$ with respect to the objective probability $P$ and $D_i^t \equiv \exp(-\rho_i t)$ the discount factor of agent $i$ between date 0 and date $t$. As previously, we consider the aggregate utility function $U$ defined as the solution of the following maximization program

$$U(e_i^t) = \max_{\sum_{i \in I} \gamma_i = \epsilon_i} \sum_{i \in I} \lambda_i E \left[ M_i^t D_i^t u(y_i^t) \right]$$

where $(\lambda_i)$ are given positive weights. Each agent is then characterized by a belief $M_i^t$, a discount factor $D_i^t$ and a weight $\lambda_i$.

We will say that the characteristics $(M_i^t, D_i^t, \lambda_i)_{i \in I}$ are independent if for almost all states of the world $\omega$, $M_i^t(\omega), D_i^t$ and $\lambda_i$ are independent as random variables on $I$. This property will be, in particular, satisfied when $I$ can be written in the form $I = J \times K \times L$ and when there exist characteristics $(M_j^i)_{j \in J}$, $(D_k^i)_{k \in K}$ and $(\lambda_{i\ell})_{\ell \in L}$ such that for $i = (j, k, \ell)$ we have $(M_i^t, D_i^t, \lambda_i) = (M_j^i, D_k^i, \lambda_{i\ell})$. Roughly speaking, this property means that there is no specific correlation between beliefs and time preferences and that the weights granted by the social planner to the individuals in the economy are independent of their time and belief characteristics.

This condition is, in particular, satisfied when beliefs and time preferences are independent and when the agents are uniformly weighted in the social welfare function. This is also the case when the agents’ weights are given by their relative wealth and when wealth, beliefs and time preferences are independent.

Assuming uniform weights is quite reasonable since there is no particular reason for the social planner to favor one agent with respect to another one. The independence of beliefs and time preference rates is more disputable. They may be positively as well as negatively correlated, the independence condition may then be analyzed as a central scenario.

---

9 More precisely, for any real valued (measurable) functions $f, g, h$ defined on the real line, we have

$$\frac{1}{n} \sum_{i \in I} f(M_i^t) g(D_i^t) h(\lambda_i) = \left( \frac{1}{n} \sum_{i \in I} f(M_i^t) \right) \left( \frac{1}{n} \sum_{i \in I} g(D_i^t) \right) \left( \frac{1}{n} \sum_{i \in I} h(\lambda_i) \right) \text{ a.e.}$$
We easily obtain the analog of Proposition 1 in our setting.

**Proposition 8** If the characteristics \((M^i_t, D^i_t, \lambda_i)_{i \in I}\) are independent, then the aggregate utility for consumption is given by

\[
U(e^*_t) = E[M_tD_tu(e^*_t)]
\]

with

\[
M_t = \left( \frac{1}{|I|} \sum_{i \in I} (M^i_t)^\eta \right)^{\frac{1}{\eta}} \quad \text{and} \quad D_t = \left( \frac{1}{|I|} \sum_{i \in I} (D^i_t)^\eta \right)^{\frac{1}{\eta}}.
\]

The representative agent belief is then given by \(M_t = \left( \frac{1}{|I|} \sum_{i \in I} (M^i_t)^\eta \right)^{\frac{1}{2}}\) and the representative agent time discount factor is given by \(D_t = \left( \frac{1}{|I|} \sum_{i \in I} (D^i_t)^\eta \right)^{\frac{1}{2}}\).

This means that all the properties already established in the previous sections on the belief of the representative agent remain valid.

The properties of the representative agent time preference rate are easy to obtain. Note that the properties of a “consensus” time preference rate when there is heterogeneity on the individual time preference rates (and not on the beliefs) has already been studied in varying contexts. Indeed, the problem of the aggregation of the utility discount rates has been studied by Reinschmidt (2002) through a certainty equivalent approach, by Gollier-Zeckhauser (2005) and Nocetti and al. (2008) through a Benthamite/Pareto optimal approach, and by Lengwiler (2005) through an equilibrium approach. All these papers adopt a deterministic setting with no divergence on the beliefs of the agents. On the contrary our aim here is to derive the properties at the aggregate level simultaneously on the beliefs and on the time preference rate (and in a quite general stochastic setting).

We know that the representative agent time discount factor is given by \(D_t = \left( \sum_{i \in I} \frac{1}{|I|} (D^i_t)^\eta \right)^{\frac{1}{\eta}}\) where \(D^i_t \equiv \exp(-\rho_it)\). We introduce the representative agent marginal time preference rate.
\( \rho_m \) as well as the representative agent average time preference rate \( \rho_a \), respectively defined as

\[
\rho_m^D (t) = -\frac{D'_t}{D_t}, \\
\rho_a^D (t) = -\frac{1}{t} \log D_t.
\]

The average discount rate corresponds to the rate which, if applied constantly for all intervening years, would yield the discount factor \( D_t \), whereas the marginal discount rate is the rate of change of the discount factor. It is easy to recover the average discount rate from the marginal discount rate since \( \rho_a (t) = \frac{1}{t} \int_0^t \rho_m (s) \, ds \).

Let us state the following properties of the average and marginal time preference rates that are similar to those underlined by Gollier and Zeckhauser (2005) and Lengwiler (2005) in similar but quite different settings.

**Proposition 9** Properties of the representative agent time preference rate

1. The representative agent average and marginal time preference rates are given by

\[
\rho_a^D (t) = -\frac{1}{t} \log \left[ \frac{1}{N} \sum_{i=1}^{N} \exp (-\eta \rho_i t) \right]^{1/\eta}, \\
\rho_m^D (t) = \sum_{i=1}^{N} \frac{\exp (-\eta \rho_i t)}{\sum_{i=1}^{N} \exp (-\eta \rho_i t)} \rho_i.
\]

2. The representative agent time preference rates are lower than the average of the time preference rates, i.e.

\[
\rho_m^D (t) < \frac{1}{N} \sum_{i=1}^{N} \rho_i \quad \text{and} \quad \rho_a^D (t) < \frac{1}{N} \sum_{i=1}^{N} \rho_i.
\]

3. “Behavioral Properties” : The representative agent time preference rates are decreasing with time. Moreover, the asymptotic discount rates are given by the lowest time preference rate, i.e. \( \lim_{t \to +\infty} \rho_a^D (t) = \lim_{t \to +\infty} \rho_m^D (t) = \inf_i (\rho_i) \). The representative agent behaves
for $t$ large enough like the most patient agent.

The expression for the marginal time preference rate in the setting of logarithmic utility functions coincides with the expression obtained by Nocetti et al. (2008) even though the approaches are different. The results on $\rho_m^D$ can be obtained as immediate extensions of those Nocetti et al. (2008).

These formulas permit explicit computations for specific distributions of the individual time preference rates. For instance, if we assume a Gamma\(^{10}\) distribution $\gamma(\alpha, \beta)$ for the $\rho_i$s we obtain

$$\rho_m^D(t) = \frac{m^2}{m + \eta v^2 t}$$

where $m$ and $v^2$ respectively denote the mean and the variance of the considered distribution.

It is immediate on this simple example that the marginal discount rate decreases with time and is hyperbolic as in Weitzman (1998, 2001). Furthermore, the speed of the decrease increases with the level of heterogeneity $v^2$ as well as with the level of risk tolerance.

It is easy to verify that the "hyperbolic" property as well as the asymptotic property remain valid for non constant time preference rates as long as these rates are nonincreasing.

Let us now analyse more in detail the impact of the choice of the distribution $f_\rho$ on the average and marginal discount rates. More precisely, the next proposition provides comparative statics results for shifts of the distribution $f_\rho$.

**Proposition 10**

1. A FSD (resp. SSD) dominated shift on the distribution $f_\rho$ of individual marginal time preference rates decreases the representative agent average time preference rate $\rho_a^D$.

2. A MLR (resp. PD) dominated shift on the distribution $f_\rho$ of individual marginal time preference rates decreases the representative agent average time preference rate $\rho_m^D$.

\(^{10}\)As mentioned in Section 2, sums should be replaced by integrals when dealing with continuous distributions. The density function of a gamma distribution $\gamma(\alpha, \beta)$ is given by $\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x)$. Its mean $m$ and its variance $v^2$ are respectively given by $m = \frac{\alpha}{\beta}$ and $v^2 = \frac{\alpha}{\beta^2}$. 22
Second Stochastic Dominance as well as Portfolio Dominance are related to a notion of risk or of dispersion while First Stochastic Dominance and Monotone Likelihood Ratio Dominance are related to notions of shifts from low values to high values. Roughly speaking, Proposition 10 introduces the right concepts of dispersion and shifts and shows that more dispersion in agents’ time preference rates as well as shifts to lower values of individual time preference rates decrease the representative agent’s time preference rate.

7 Interpretation, Applications and Discussion

We have seen that starting from a standard model with optimistic as well as pessimistic vNM and exponential discounting agents, we obtain at the representative agent level properties such as an inverse S-shaped distribution probability weighting function and hyperbolic discounting, that are in line with recent empirical and experimental results. A possible interpretation of such a result is to consider that each individual subject to experiments behaves as a group of individuals at the equilibrium. This provides us with a possible representation of the brain as an organization with a social planner and heterogeneous doers. This also provides us with a new tool to analyze the equilibrium properties in models with "behavioral" agents. It suffices to replace each "behavioral" agent by a family of "classical" heterogeneous agents and to embed our analysis in the analysis of equilibria with heterogeneous beliefs and heterogeneous time preference rates as in Jouini and Napp (2007).

In this section we discuss these interpretations and applications. We also analyze how our model can be extended to a dynamic framework (more than 2 dates).

7.1 The brain as a central planner and heterogeneous doers

Our results suggest a model for the brain where a central planner (the cortex) relies on heterogeneous doers in order to evaluate prospects. Some doers are overoptimistic while others are overpessimistic. Similarly, some doers are impatient while others are more patient. Such an
approach is in the same spirit as Brocas and Carillo (2008) where the authors divide the brain into two systems: an impulsive/myopic one and a cognitive/forward-looking one. However, while Brocas and Carillo (2008) model mainly relies on information asymmetries and principal-agent models, our model relies on decentralization and optimal allocation approaches. Such a decomposition of the brain into different systems that are possibly in conflict are based on recent neuroscience and psychology evidences related to intrapersonal tensions (see e.g. Brocas and Carillo, 2008, for an extensive discussion of this literature): temporal horizon conflicts, information conflicts (that may lead to information asymmetries as in Brocas and Carillo, 2008, but also to information diversity and beliefs heterogeneity as in our model) or utility evaluation conflicts. In economics, such decompositions have been first considered by Thaler and Shefrin (1981) and Shefrin and Thaler (1988).

Let us describe more in detail how our heterogeneous agents model may be applied to represent the brain. We assume that the brain relies on separate systems in order to evaluate a given prospect. These systems are specialized in the sense that each system evaluates a given "dimension" of the prospect under consideration. For instance, one system may evaluate the positive part of the prospect while another system evaluates the negative part of the same prospect. In such a case gains and losses will be treated differently as in Prospect Theory. The prospect under consideration is then decomposed and each part of the decomposition is attributed for evaluation to the appropriate system. Such an evaluation of different "dimensions" by different systems is consistent with the recent findings of Boassaerts et al. (2006) where the authors show that risk and return are evaluated separately. We assume then that each system is characterized by a given belief (relative weights of the different states of the world). Such beliefs heterogeneity might result from information conflicts or from noise in the transmission of information to the different systems or from biases in the treatment of this information (some systems are overoptimistic while others are overpessimistic). Each system also has its own time preference rate that corresponds to its own level of impatience. We finally assume that the decomposition of
the prospect under consideration and the distribution of the different components among the systems is made endogeneously; each marginal allocation is given to the system that valuates it more leading to a Pareto optimum\textsuperscript{11}. Such a Pareto optimum might result from a centralized allocation process organized by a central planner or from a tatonnement process where each decomposition leads to an overall evaluation (the sum of the individual evaluations) and where only the best overall evaluation (after a given number of trials) is kept.

Let us consider the simplest case for such a model. The brain faces a lottery whose average payoff is given by $\mu$. This information is passed on two separate systems. Due to noise, the first system perceives $\mu + \varepsilon$ and the second one perceives $\mu - \varepsilon$ (on average there is no specific reasons for the average perceived signal to be biased). The lottery is then modeled in each system by a normal distribution centered around the perceived signal and with a given variance $\sigma^2$. The "optimistic" system is more interested by the right side of the distribution and the "pessimistic" system is more interested by the left side of the distribution. A Pareto optima decomposition leads then to an overall (representative agent) evaluation that corresponds to the valuation that would be provided by a "behavioral" agent. The level of discriminability is then directly related to the level of noise as shown in Proposition 7. The level of attractiveness would be associated to the level of systematic bias (if any) as shown in Proposition 6.

### 7.2 Equilibrium properties of "behavioral" economies

Recent developments in decision theory permit to better reflect individual behavior. However they are difficult to handle in an equilibrium framework. Our results imply that we can interpret each agent with behavioral preferences and decreasing discount rates as a collection of agents with heterogeneous standard vNM preferences and exponential discounting.

More precisely, let us consider a model with $n$ agents with behavioral probability weighting

\textsuperscript{11}For instance, if one system has a belief whose support is $\mathbb{R}_+$ while a second system has a belief whose support is $\mathbb{R}_-$, then at the Pareto optimum, the first system will treat the positive part of the prospect while the second system will treat its negative part.

25
functions and hyperbolic discount rates. The probability weighting function and the discounting function of a given agent can be obtained as the representative agent weighting and discounting functions in a model with four agents that have heterogeneous standard vNM preferences and exponential discounting. Indeed, as seen in the previous sections it is possible to construct $M^1$ and $M^2$ in order to fit the density $M$ of the agent under consideration through the formula $M = \left( \sum_{i=1}^{2} \gamma_i (M^i)^\eta \right)^{\frac{1}{\eta}}$ and to find $D^1$ and $D^2$ in order to fit the discounting function $D$ of the agent under consideration through the formula $D_t = \left( \sum_{i=1}^{2} \gamma_i (D^i_t)^\eta \right)^{\frac{1}{\eta}}$ and we have then, by construction

$$E[MD_t u(x)] = \max_{\sum_{i,j=1}^{2} x_{i,j} = x} \sum_{i,j=1}^{2} E(M^i D^j_t u(x_{i,j})).$$

The valuation function of each agent then corresponds to the social valuation function in a model with 4 agents. Such a family $(M^i, D^j)_{i,j=1,2}$ may be constructed for each agent and we will denote it by $(M^{n,i}, D^{n,j})_{i,j=1,2}$ when applied to agent $n$. In order to recover a Pareto Optimum $(x_k)_{k=1,...,n}$ in the initial economy with $n$ agents, it suffices to look for a Pareto Optimum $(x_{i,j}^k)_{i,j=1,2, k=1,...,n}$ in the 4$n$ agents economy and to take $x_k = \sum_{i,j=1}^{2} x_{i,j}^k$. Indeed

$$\max_{\sum_{i,j=1}^{2} x_{i,j}^k = x} \sum_{k=1}^{n} E(M^k D^j_t u(x_k)) = \max_{\sum_{i,j=1}^{2} x_{i,j}^k = x} \sum_{i,j=1, k=1,...,n} E(M^{k,i} D^{j,k}_t u(x_{i,j}^k))$$

and the solutions of the first maximization problem may be recovered from the solutions of the second one through the formulas above.

Pareto Optima in "behavioral" models correspond then to Pareto Optima in heterogeneous beliefs and heterogeneous time preference rates models (with a larger number of agents). These models share then the same properties at least at the aggregate level (state price density, yield curve, etc.). Models with heterogeneous beliefs have been studied for instance by Jouini and Napp (2007) and lead to tractable results. Many applications may follow. For instance, if in a behavioral model the agents have probability weighting functions that intersect the diagonal
below 0.5 (which is the typical situation) then each agent corresponds to a combination of an optimistic and of a pessimistic agent where the pessimistic agent is more pessimistic than the optimistic is optimistic. This leads to a pessimistic bias at the aggregate level and then to a higher risk premium (see Jouini and Napp, 2007).

7.3 Dynamic setting

All the results stated in this paper were in a two dates setting even though we considered different possible time horizons for the second date. If we want to exploit better the similarities between "behavioral" models and heterogeneous agents models that are described in the previous section, it is useful to understand how our model might be embedded in a dynamic setting and, in particular, to understand how the representative agent probability weighting function is affected in such a setting.

It can be seen that, if \( e_t^i \sim \mathcal{N}(\mu_t, \sigma^2 t) \) as in a diffusion framework and if the subjective distributions are of the form \( \mathcal{N}(\mu_t, \sigma^2 t) \) with \( \mu_t = \mu t + \delta_i(t)\sigma \sqrt{t} \), the probability weighting function is time independent if and only if \( \delta_i(t) \) is constant. This means that agents’ deviation from the objective mean \( \mu t \) is constant in terms of standard deviation \( \mu_t = \mu t + \delta_i\sigma \sqrt{t} \). However, agents’ deviation from the objective drift \( \mu \) decreases with the horizon, \( \mu_i = \mu + \delta_i \frac{\sigma}{\sqrt{t}} \).

Let us now analyse how mean and variance evolve in a dynamic setting and, for this purpose, let us consider a wealth process \( e_t^i \) that follows a geometric Brownian motion with drift \( \mu + \frac{1}{2}\sigma^2 \) and volatility \( \sigma \). The distribution of \( e_t^i \) is then lognormal and we have \( \log e_t^i \sim \mathcal{N}(\mu_t, \sigma_t^2) \) with \( \mu_t = \mu t \) and \( \sigma_t^2 = \sigma^2 t \). By Girsanov Theorem, the subjective distribution of \( e_t^i \) from agent \( i \) point of view is necessarily of the form \( \log e_t^i \sim \mathcal{Q}_i \mathcal{N}(\mu_i^t, \sigma_i^2) \). Let us assume that Group 1 overestimates the drift (average return per unit of time) by \( \delta \) times the volatility (standard deviation of the return per unit of time) while Group 2 underestimates the drift by the same quantity. We have then \( \mu_i^t = \mu t \pm \delta \sigma t \). As far as the excess variance is concerned, we have \( \text{Var}_i(\mu_i^t) = \sigma^2 t^2 \delta^2 \) which dominates \( \text{Var}^P[\log e_i^t] = \sigma^2_t = \sigma^2 t \), for \( t \) large enough. The long
term volatility, from the representative agent point of view, is then arbitrarily large when the horizon becomes longer. If we now assume that $\mu^i_t = \mu t \pm \delta \sigma \sqrt{t}$ (agent’s i deviation from the objective expected logarithmic return $\mu_t = \mu t$ is equal to $\delta$ times the objective standard deviation of $\log e^*_t$) we have $\text{Var}^M [\log e^*_t] = \sigma^2 (1 + \delta^2) t$. The long term volatility, from the representative agent point of view, is then higher than the objective volatility and the difference between the subjective and the objective volatilities is constant and given by $\sigma^2 \delta^2$ which is a measure of the divergence of beliefs among the agents. In both cases, the mean remains the same in the unbiased setting and the variance per unit of time is higher.

We recover then in the dynamic setting the same qualitative properties as in the static setting. However the scale of the different phenomena strongly depend on how beliefs evolve through time, which may be subject to future developments.

References


Reinschmidt, K.F., 2002. Aggregate Social Discount Rate derived from individual discount


Proof of Proposition 1

At the Pareto optimum, we have

\[ \lambda_i M_i^t u'(y_i^t) = q_t \]

for some random variable \( q_t \). It follows that

\[ y_i^t = \left[ \frac{q_t}{\lambda_i M_i^t} \right]^{-\eta} \]

hence

\[ e_i^* = \sum_{i \in I} \left[ \frac{q_t}{\lambda_i M_i^t} \right]^{-\eta} = q_t^{-\eta} \sum_{i \in I} \left[ \frac{1}{\lambda_i M_i^t} \right]^{-\eta} \]

and

\[ y_i^t = e_i^* \frac{[\lambda_i M_i^t]^\eta}{\sum_{i \in I} [\lambda_i M_i^t]^\eta}. \]

We have then

\[
\begin{align*}
\sum_{i \in I} \lambda_i E \left[ M_i^t u(y_i^t) \right] &= \sum_{i \in I} \lambda_i E \left[ M_i^t \frac{[\lambda_i M_i^t]^{\eta-1}}{(\sum_{i \in I} [\lambda_i M_i^t]^\eta)^{1-\frac{1}{\eta}}} u(e_i^*) \right] \\
&= E \left[ \frac{\sum_{i \in I} [\lambda_i M_i^t]^\eta}{(\sum_{i \in I} [\lambda_i M_i^t]^\eta)^{1-\frac{1}{\eta}}} u(e_i^*) \right] \\
&= E \left[ \left( \sum_{i \in I} [\lambda_i M_i^t]^\eta \right)^{1/\eta} u(e_i^*) \right]
\end{align*}
\]

and \( U(e^*) = E [Mu(e^*)] \) with \( M = \left[ \sum_{i \in I} \lambda_i^\eta (M_i^t)^\eta \right]^{1/\eta}. \)

Proof of Proposition 2
We have

\[ E[M_i h(e_i)] = E \left[ \left( \sum_{i \in I} \gamma_i (M)^\eta \right)^{1/\eta} h(e_i) \right] \]

\[ = E \left[ \left( \sum_{i \in I} \gamma_i (f_i(e_i^*))^{\eta} \right)^{1/\eta} h(e_i) \right] \]

\[ = E \left[ \frac{\left( \sum_{i \in I} \gamma_i (f_i(e_i^*))^{\eta} \right)^{1/\eta}}{f(e_i^*)} h(e_i) \right] \]

\[ = \int \frac{\left( \sum_{i \in I} \gamma_i (f_i(x))^{\eta} \right)^{1/\eta}}{f(x)} h(x) f(x) dx \]

\[ = \int \frac{\left( \sum_{i \in I} \gamma_i (f_i(x))^{\eta} \right)^{1/\eta}}{f(x)} h(x) dx \]

hence \( f^M = \left( \sum_{i \in I} \gamma_i (f_i)^\eta \right)^{1/\eta} \).

\[ \text{Proof of Proposition 3} \]

1. Since a mixture of Gaussian distributions is not Gaussian, the first part is immediate. For the second part, we have

\[ (f^{\log})^\eta = \frac{1}{2} \left( f_1^{\log} \right)^\eta + \frac{1}{2} \left( f_2^{\log} \right)^\eta = \frac{1}{2} \exp \left( -\frac{\eta (x - \mu_1)^2}{2\sigma^2} \right) \]

\[ + \frac{1}{2} \exp \left( -\frac{\eta (x - \mu_2)^2}{2\sigma^2} \right) . \]

This function has either two maxima that are symmetric with respect to \( \frac{\mu_1 + \mu_2}{2} \) or only one maximum at \( \frac{\mu_1 + \mu_2}{2} \). In the first case \( \frac{\mu_1 + \mu_2}{2} \) would be a local minimum. It suffices then to analyse the sign of the second derivative of \( (f^{\log})^\eta \) at \( \frac{\mu_1 + \mu_2}{2} \). We obtain that the distribution is bimodal for \( |\mu_1 - \mu_2| > 2\sigma/\sqrt{\eta} \) and unimodal for \( |\mu_1 - \mu_2| \leq 2\sigma/\sqrt{\eta} \).
2. For \( \eta = 1 \), we have \( E^Q \log e^* = \frac{1}{2} E^Q \log e^* + \frac{1}{2} E^Q \log e^* = \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 \). We have

\[
\text{Var}^Q \log e^* = E^Q (\log e^*)^2 - E^Q (\log e^*)^2 = E^Q (\log e^*)^2 - \left( E^Q \log e^* \right)^2 = \frac{1}{2} \left( \text{Var}^Q \log e^* + \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 \right)
\]

\[
+ \frac{1}{2} \left( \text{Var}^Q \log e^* + \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 \right) - \left( \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 \right)^2 = \text{Var}^P \log e^* + \frac{1}{4} (\mu_1 - \mu_2)^2 = \sigma^2 + \text{Var}_i (\mu_i)
\]

We have

\[
E^Q (\log e^* - E^Q (\log e^*))^4 = E^Q (\log e^*)^4 - 4 E^Q (\log e^*)^3 E^Q (\log e^*) + 6 E^Q (\log e^*)^2 E^M (\log e^*) - 4 E^Q (\log e^*) E^Q (\log e^*)^3 + E^Q (\log e^*)^4.
\]

We easily get that \( E^{Q_i} (\log e^*)^2 = \sigma^2 + \mu_i^2 \), \( E^{Q_i} (\log e^*)^3 = \mu_i \left( \mu_i^2 + \sigma^2 \right) \), and \( E^{Q_i} (\log e^*)^4 = \mu_i^4 + 6 \mu_i^2 \sigma^2 + 3 \sigma^4 \).

Suppose that \( \bar{\mu} \equiv \mu_1 + \mu_2 = 0 \). We get that \( E^Q (\log e^* - E^Q (\log e^*))^4 = \frac{1}{2} (\mu_1^4 + \mu_2^4) + 3 (\mu_1^2 + \mu_2^2) \sigma^2 + 3 \sigma^4 \) and \( E^Q (\log e^* - E^Q (\log e^*))^2 = \sigma^2 + \frac{1}{2} (\mu_1^2 + \mu_2^2) \) hence

\[
\frac{E^Q (\log e^* - E^Q (\log e^*))^4}{E^Q (\log e^* - E^Q (\log e^*))^2} = \frac{\mu_1^4 + 6 \mu_i^2 \sigma^2 + 3 \sigma^4}{\mu_1^4 + 2 \mu_i^2 \sigma^2 + \sigma^4} = \frac{3 - \frac{2 \mu_1^4}{\mu_1^4 + 2 \mu_i^2 \sigma^2 + \sigma^4}}{\frac{2 \mu_1^4}{\mu_1^4 + 2 \mu_i^2 \sigma^2 + \sigma^4}} < 3.
\]

For general \( \bar{\mu} \), it suffices to translate uniformly all the considered distributions to obtain the result.

3. The ratio between the density of \( \log e^* \) under \( Q \) and the density of \( \log e^* \) under \( P \) is given
by

\[
\frac{f^M \log}{f \log}(x) = \left( \frac{1}{2} \exp \left( \frac{-2(x - \mu)(\mu - \mu_1) + \mu^2 - \mu_1^2}{2\sigma^2} \right) + \frac{1}{2} \exp \left( \frac{-2x(\mu - \mu_2) + \mu^2 - \mu_2^2}{2\sigma^2} \right) \right)^{\frac{1}{2}}
\]

which is clearly symmetric with respect to \( \mu \), decreasing before \( \mu \) and increasing after \( \mu \).

Moreover, since the distributions of \( \log e^* \) under \( Q \) and under \( P \) are both symmetric with respect to \( \mu \), we have \( E^Q [\log e^*] = E^P [\log e^*] = \mu \). These properties give \( \text{Var}^Q [\log e^*] > \text{Var}^P [\log e^*] \) (see Jouini and Napp, 2008).

4. For general \((\mu_1)\), \( f^M \log \) is symmetric with respect to \( \frac{\mu_1 + \mu_2}{2} \) which gives \( E^Q [\log e^*] = \frac{\mu_1 + \mu_2}{2} \).

Furthermore, we may apply the same reasoning as in 3. to compare the distribution of \( \log e^* \) under \( Q \) with the distribution whose density is given by \( \frac{1}{\sqrt{2\pi}\sigma} \exp \left( -\frac{(x - \frac{\mu_1 + \mu_2}{2})^2}{2\sigma^2} \right) \). We then have \( \text{Var}^Q [\log e^*] > \sigma^2 = \text{Var}^P [\log e^*] \).

5. For two different values \( \eta \) and \( \eta' \) of the risk tolerance parameter, it suffices to consider \( \frac{f^M \log}{f^M \log'} \) and to apply the same reasoning as in 3.

\[\blacksquare\]

**Proof of Proposition 4**

1. If \( \lim_{\infty} \frac{f_{\text{opt}}}{f} = \lim_{-\infty} \frac{f_{\text{pass}}}{f} = \infty \) and \( \lim_{-\infty} \frac{f_{\text{opt}}}{f} = \lim_{-\infty} \frac{f_{\text{pass}}}{f} = 0 \) then the representative agent density function is such that \( \lim_{-\infty} \frac{f^M}{f} = \lim_{-\infty} \frac{f^M}{f} = \infty \) and \( \frac{f^M}{f} \) can not be monotone.

2. This is immediate according to \( \lim_{-\infty} \frac{f_{\text{opt}}}{f} = \lim_{-\infty} \frac{f_{\text{pass}}}{f} = \infty \).

3. It suffices to remark that \( Q(x \leq \underline{x}) = \int_{-\infty}^{\underline{x}} f_M(s) \, ds \) and to apply 2. to conclude that \( Q(x \geq \underline{x}) \geq P(x \geq \underline{x}) \) for \( \underline{x} \leq x_{\text{inf}} \). The same applies for \( \overline{x} \geq x_{\text{sup}} \).

4. It suffices to remark that \( f^M = f_{\text{opt}}^\max \left( \gamma_{\text{opt}}^\max + \sum_{i=1, \ldots, N} \left( \frac{f_i}{f_{\text{opt}}^i} \right)^{\eta} \right)^{1/\eta} \). If \( \frac{f_i}{f_{\text{opt}}^i} \) is non-increasing for all \( i \) then \( \left( \gamma_{\text{opt}}^\max + \sum_{i=1, \ldots, N} \left( \frac{f_i}{f_{\text{opt}}^i} \right)^{1/\eta} \right)^{1/\eta} \) is bounded away from 0 and \( \infty \) in
the neighborhood of $\infty$ and we have $f^M \sim f_{\text{opt}}^{\max}$. The result at the neighborhood of $-\infty$ is obtained similarly.

\section*{Proof of Proposition 5}

1. Let $g$ be given by $\int_{-\infty}^{\infty} f_M(x)dx = g \left[ \int_{-\infty}^{\infty} f(x)dx \right]$. We have $f_M(x) = g' \left[ \int_{-\infty}^{\infty} f(x)dx \right] f(u)$ and $g' \left[ \int_{-\infty}^{\infty} f(x)dx \right] = \frac{f_M}{f} (u)$. We also have $-f(u)g'' \left[ \int_{-\infty}^{\infty} f(x)dx \right] = \left( \frac{f_M}{f} \right)' (u)$ which gives that the concavity of $g$ is governed by the sign of $\left( \frac{f_M}{f} \right)'$. Remark that $\left( \frac{f_M}{f} \right)'$ is negative in a neighborhood of $-\infty$ and $g''$ is positive and $g$ is convex in a neighborhood of 1. Similarly, we have that $\left( \frac{f_M}{f} \right)'$ is positive in a neighborhood of $\infty$ and that $g''$ is negative and $g$ is concave in a neighborhood of 1. Finally, $\left( \frac{f_M}{f} \right)'$ is a combination of exponentials where the decreasing exponentials have a negative weight and the increasing exponentials have a positive weight. The function $\left( \frac{f_M}{f} \right)'$ is then increasing. The function $g$ is then inverse S-shaped: concave then convex.

2. Let $g$ be given by $\int_{-\infty}^{\infty} f_M(x)dx = g \left[ \int_{-\infty}^{\infty} f(x)dx \right]$. We have $f_M(u) = g' \left[ \int_{-\infty}^{\infty} f(x)dx \right] f(u)$ and $g' \left[ \int_{-\infty}^{\infty} f(x)dx \right] = \frac{f_M}{f} (u)$. In particular, we have $g'(0) = \frac{f_M}{f} (\infty) = \infty$. If $g''(0)$ is well defined, we have $g''(0) < 0$ and hence $g''(x) < 0$ in a neighborhood of 0. The probability weighting function is then concave for small probabilities. The result in the neighborhood of 1 is obtained similarly.

\section*{Proof of Proposition 6}

1. Let us consider a distribution of the means that is described by a density function $h$. The associated representative agent cumulative distribution function is given by

$$
\frac{1}{\sqrt{2\pi\sigma^2}} \int dh(\mu) \int_{-\infty}^{x} \exp -\frac{(s-\mu)^2}{2\sigma^2} ds.
$$

Since the function $\mu \rightarrow \int_{-\infty}^{x} \exp -\frac{(s-\mu)^2}{2\sigma^2} ds$ is decreasing a FSD shift of $h$ decreases the value of $\int dh(\mu) \int_{-\infty}^{x} \exp -\frac{(s-\mu)^2}{2\sigma^2} ds$ and leads then to a FSD dominating distribution function for the representative agent.
2. Let us consider a distribution \((\gamma_i')\) and a FSD dominated shift \((\gamma_i)\). We want to prove that \(\sum \gamma'_i F_i \geq \sum \gamma_i F_i\). For a given \(x\), letting \(x_i\) denote the quantity \(F_i(x)\), it suffices to prove that \(\sum \gamma'_i x_i \geq \sum \gamma_i x_i\) for a nondecreasing family \((x_i)_{i \in I}\) which is true since \((\gamma_i')\) dominates \((\gamma_i)\) in the sense of the FSD.

3. Let us consider a distribution \((\gamma_i')\) and a MLR dominated shift \((\gamma_i)\). It suffices to prove that \((P_{0 i} f_i) \geq (P_{i} f_i)\) is increasing or that \(\frac{\sum \gamma'_i F_i}{\sum \gamma_i F_i}\) is increasing with \(F_i = f_i^0\). Without any loss of generality, we may assume that all the considered functions are differentiable and let us consider the derivative of \(\frac{\sum \gamma'_i F_i}{\sum \gamma_i F_i}\)

\[
\left( \frac{\sum \gamma'_i F_i}{\sum \gamma_i F_i} \right)' = \left( \frac{\sum \gamma'_i F_i (\sum \gamma_i F_i) - (\sum \gamma'_i F_i) (\sum \gamma_i F_i')}{(\sum \gamma_i F_i)^2} \right) \\
= \frac{\sum f_i \gamma_i f_j (\frac{\gamma'_i}{\gamma_i} - \frac{\gamma'_j}{\gamma_j}) (F'_i F_j - F_i F'_j)}{(\sum \gamma_i F_i)^2}.
\]

Remark that for \(f_i \geq f_j\) we have \(F_i \geq F_j\) and then \(F'_i F_j - F_i F'_j \geq 0\). Furthermore, for \(f_i \geq f_j\) we also have \(\frac{\gamma'_i}{\gamma_i} - \frac{\gamma'_j}{\gamma_j} \geq 0\) which leads to the conclusion.

**Proof of Proposition 7**

It is immediate that \(\mu_1, \mu_2,\) and \(\mu_3\) can be written on the form \(\mu_2 - h, \mu_2, \mu_2 + h\) for some \(h > 0\). For the distribution of individual characteristics \((\gamma_i)\), the representative agent distribution function is given by \(\frac{1}{\sqrt{2\pi\sigma^2}} \sum \gamma_i \int_{-\infty}^{x} \exp (-\frac{(s-\mu_i)^2}{2\sigma^2}) ds\). The symmetric mean preserving spread induces a modification of this distribution that is positively proportional to \(\frac{1}{\sqrt{2\pi\sigma^2}} \left( \int_{-\infty}^{x} \exp \left( -\frac{(s-\mu_2+h)^2}{2\sigma^2} \right) - 2 \exp \left( -\frac{(s-\mu_2)^2}{2\sigma^2} \right) + \exp \left( -\frac{(s-\mu_2-h)^2}{2\sigma^2} \right) \right) ds\). Simple computations permit to show that this modification is positive for \(x \leq \mu_2\) and negative for \(x \geq \mu_2\). A symmetric mean preserving spread leads then to a distribution function that is above the original one below a given threshold and below the original one above the threshold. We have then an increase of the level of discriminating ability.

**Proof of Proposition 8**
Replacing $M^i$ by $M^iD^i$ in the proof of Proposition 1, we easily get that

$$U(e^0) = \left[ \sum_{i \in I} [\lambda_i M_i^i D_i^i]^{\eta} \right]^{1/\eta}.$$  

Now, if the characteristics $(\lambda_i, M_i^i, D_i^i)$ are independent, then

$$\left[ \sum_{i \in I} [\lambda_i M_i^i D_i^i]^{\eta} \right]^{1/\eta} = \left[ \left( \frac{1}{|I|} \sum_{i \in I} (M_i^i)^\eta \right) \right]^{1/\eta} \left[ \left( \frac{1}{|I|} \sum_{i \in I} (D_i^i)^\eta \right) \right]^{1/\eta}$$

and

$$\sum_{i \in I} \lambda_i E \left[ M_i^i D_i^i u(y_i^i) \right] = E \left[ \left( \frac{1}{|I|} \sum_{i \in I} (M_i^i)^\eta \right) \left( \frac{1}{|I|} \sum_{i \in I} (D_i^i)^\eta \right) u(e_i^0) \right].$$

Proof of Proposition 9

We prove the proposition for $\rho^D_m$ since it is easy to check that all the derived properties are inherited by $\rho^D_a(t) = \frac{1}{t} \int_0^t \rho^D_m(s) ds$.

1. Immediate.

2. The representative agent time preference rate $\rho^D_m(t) = \sum_{i=1}^N \frac{\exp(-\eta \rho_i t)}{\sum_{i=1}^N \exp(-\eta \rho_i t)} \rho_i$ is an average of the $\rho_i$s with weights that decrease with $\rho_i$. Such an average is smaller than the equally weighted average.

3. Denote $\theta_i = \exp(-\eta \rho_i t)$. We have $\frac{d \rho^D_m(t)}{dt} = -\left( \frac{\sum_{i \in I} \theta_i \rho_i^2}{\sum_{i \in I} \theta_i} - \left( \frac{\sum_{i \in I} \theta_i \rho_i}{\sum_{i \in I} \theta_i} \right)^2 \right)$ which is negative.

We have $\rho^D_m(t) = \frac{\rho_{inf} + \sum_{i \neq inf} \exp(-\eta(\rho_i - \rho_{inf}) t) \rho_i}{1 + \sum_{i \neq inf} \exp(-\eta(\rho_i - \rho_{inf}) t)}$ and $\exp(-\eta(\rho_i - \rho_{inf}) t) \rho_i \to _\infty 0$ we have then $\rho^D_m(t) \to \rho_{inf}$.

Proof of Proposition 10
The proof of 1. is inspired from Jouini and Napp (2008) and the proof of 2. is inspired from Nocetti et al. (2008).

1. We have

\[ \rho_a^D (t) \equiv - \frac{1}{\eta t} \ln E [\exp (-\eta \rho t)] \]

where \( E \) is the expectation operator associated with the distribution of \( (\rho_i) \). For a given \( t \), the function \( \rho \to \exp (-\eta \rho t) \) is decreasing (and convex) and, by definition, a FSD (resp. SSD) shift in the distribution of \( (\rho_i) \) decreases the value of \( E [\exp (-\eta \rho t)] \) and increases \( \rho_a^D (t) \).

2. We have then

\[ \rho_m^D (t) = \frac{E [\rho \exp (-\eta \rho t)]}{E [\exp (-\eta \rho t)]} . \]

where \( E \) is the expectation operator associated with the distribution of \( (\rho_i) \).

Let us now consider \( P^1 \) and \( P^2 \), two distributions such that \( P^2 \succeq_{MLR} P^1 \). By definition, the density \( \phi = \frac{dP^2}{dP^1} \) is nondecreasing in \( \rho \) (in other words \( i \to \phi^i \) and \( i \to \rho_i \) are comonotonic). We have then,

\[ \frac{E_{P^2} [\rho \exp (-\eta \rho t)]}{E_{P^2} [\exp (-\eta \rho t)]} = \frac{E_{P^1} [\phi \rho \exp (-\eta \rho t)]}{E_{P^1} [\phi \exp (-\eta \rho t)]} = \frac{E_{Q} [\phi \rho]}{E_{Q} [\phi]}. \]

where \( Q \) is defined by a density with respect to \( P^1 \) equal (up to a constant) to \( \exp (-\eta \rho t) \). Since \( \phi \) is nondecreasing in \( \rho \), we have

\[ E_{Q} [\phi \rho] \geq E_{Q} [\phi] E_{Q} [\rho], \]

hence

\[ \frac{E_{P^2} [\rho \exp (-\eta \rho t)]}{E_{P^2} [\exp (-\eta \rho t)]} \geq E_{Q} [\rho], \]

\[ \geq \frac{E_{P^1} [\rho \exp (-\eta \rho t)]}{E_{P^1} [\exp (-\eta \rho t)]}. \]

Let us assume now that \( P^2 \succeq_{PD} P^1 \) and let us consider \( \rho_m^D, P^2 (t) \) and \( \rho_m^D, P^1 (t) \) the
associated representative agent time preference rates. We have

\[ \rho_{m}^{D,P_2}(t) = \frac{E^{P_2}[\rho \exp(-\eta \rho t)]}{E^{P_2}[\exp(-\eta \rho t)]} \]

and then

\[ E^{P_2}[u'(\rho)(\rho - \rho_{m}^{D,P_2})] = 0 \] with \( u(\rho) = -\exp(-\eta \rho t) \). By definition, this implies

\[ E^{P_1}[u'(\rho)(\rho - \rho_{m}^{D,P_2})] \leq 0 \] hence \( \rho_{m}^{D,P_2} \geq \rho_{m}^{D,P_1} \).
Figure 1: In this figure, we have represented in black the consensus belief in a log-utility agents setting. A proportion of 47% of the agents believe that $\log e \sim \mathcal{N}(0, 1)$ and the remaining 53% believe that $\log e \sim \mathcal{N}(2.5, 1)$. The beliefs of these two categories of agents are represented in grey.

Figure 2: In this figure we represent the consensus belief for 3 different levels of risk aversion. We assume that a proportion of 47% of the agents believe that $\log e \sim \mathcal{N}(0, 1)$ and the remaining 53% believe that $\log e \sim \mathcal{N}(2.5, 1)$. The upper curve corresponds to $\eta = 2$, the lower curve to $\eta = 0.8$ and the middle curve to $\eta = 1$. An increase of $\eta$ increases the distance between the peaks and their size.

Figure 3: In this figure we represent in black the representative agent probability weighting function in a model with two logarithmic utility agents. One of them overestimates the objective mean by one standard deviation and the other one underestimates it by one standard deviation. We also represent in grey the individual probability weighting functions (the concave one corresponds to the optimistic agent).
Figure 4: In this figure we represent Prelec’s function $\exp(-(-\ln p)\gamma)$ with $\gamma = 0.73$ that corresponds to a standard specification. We also represent the probability weighting function corresponding to a model with two log-utility agents. The first one underestimates the objective average by 120% of the standard deviation and has a weight of 30%. The second one overestimates the objective average by 60% of the standard deviation and has a weight of 70%.

Figure 5: In this figure we represent the probability weighting function of the representative agent in a model with logarithmic utility agents. In the upper curve the optimistic and the pessimistic agents are equally weighted. In the lower curve, the pessimistic agents have a 60% weight and the optimistic ones have a 40% weight. Attractiveness decreases with the weight granted to the pessimistic agents.

Figure 6: The probability weighting function for different levels of divergence of belief. Both agents agree on a normal distribution but one of them overestimates the objective mean by $\delta$ times the standard deviation while the other one underestimates it by $\delta$ times the standard deviation. The value of $\delta$ ranges from 0 to 2. The discriminability decreases with $\delta$ (in other words the curvature increases with $\delta$).