Value of the future: Discounting in random environments

J. Doyne Farmer,[1,2,*] John Geanakoplos,[2,3,†] Jaume Masoliver,[4,‡] Miquel Montero,[4,§] and Josep Perelló[4,‖]

[1]*Institute for New Economic Thinking at the Oxford Martin School and Mathematical Institute, University of Oxford, Eagle House, Walton Well Rd. Oxford OX2 6ED, United Kingdom*

[2]*Santa Fe Institute, Santa Fe, New Mexico*

[3]*Department of Economics, Yale University, New Haven, Connecticut, USA*

[4]*Departament de Física Fonamental, Universitat de Barcelona, Diagonal, 645 08025 Barcelona, Spain*

(Received 14 October 2014; published 28 May 2015)

We analyze how to value future costs and benefits when they must be discounted relative to the present. We introduce the subject for the nonspecialist and take into account the randomness of the economic evolution by studying the discount function of three widely used processes for the dynamics of interest rates: Ornstein-Uhlenbeck, Feller, and log-normal. Besides obtaining exact expressions for the discount function and simple asymptotic approximations, we show that historical average interest rates overestimate long-run discount rates and that this effect can be large. In other words, long-run discount rates should be substantially less than the average rate observed in the past, otherwise any cost-benefit calculation would be biased in favor of the present and against interventions that may protect the future.

DOI: 10.1103/PhysRevE.91.052816 PACS number(s): 89.65.Gh, 88.05.Lg, 05.40.−a

I. INTRODUCTION

How can we value the future? In economics the answer to this question is given by a process called "discounting," which weights the future relative to the present. The weighting procedure is carried through a discount function which usually takes the simple form of a decreasing exponential [1]. Indeed, under a steady rate of interest r, a dollar invested today will yield e^{rt} at time t. In other words, a dollar in any future time t is worth e^{-rt} today. This simple example shows the great importance of discounting not only in finance but also in long-run environmental planning [2]. Thus an environmental problem that costs X to fix at time t is only worth an investment of $e^{-rt}X$ now. Since r tends to be substantial, any benefit at some distant time t would justify a negligible investment today. Letting interest rates be a proxy for economic growth, a different version of the same argument is that the technologies of the future will be so powerful that they will dwarf anything we can achieve with present-day technologies. Thus it is more effective to follow policies that foster economic growth than to try to combat global warming now [2–4].

The choice of a discount rate has, therefore, enormous consequences for long-run environmental planning [2,3]. For example, in a highly influential report on climate change commissioned by the UK government, Stern [4] uses a discounting rate of 1.4%, which on a 100-year horizon implies a present value of 25% (meaning the future is worth 25% as much as the present). In contrast, Nordhaus [5] argues for a discount rate of 4%, which implies a present value of 2%, and at other times [6] has advocated rates as high as 6%, which implies a present value of 0.3%. The choice of discount rate is perhaps the biggest factor influencing the debate on the urgency of the response to global warming and this issue is far from being settled.

For environmental problems normative approaches to choosing discount rates are based on ethical grounds [7,8] and assumptions about economic growth. They also depend on arguments involving the maximization of utility functions that are chosen for mathematical convenience [9]. Economists present a variety of reasons for discounting, including impatience, economic growth, and declining marginal utility; all of them are embedded in the Ramsey formula [10], which forms the basis for standard approaches to discounting [3].

However, rates are uncertain and it is not realistic to represent discounting by a deterministic function of time such as the decreasing exponential with a fixed rate and, therefore, some kind of average over all interest rate paths must be taken. This problem is particularly severe for environmental problems, where in problems such as global warming one must consider costs and benefits 100 or more years in the future. It also occurs in finance, where discounting times are typically 30 years or less, where it has long been recognized that interest rates must be modeled as random processes [11–16].

A more positive approach to discounting consists in figuring out how the market trades off present consumption for future consumption. For the near future one can readily find the corresponding market interest rate for money, and by making assumptions about likely inflation one can infer the market discount rate for real consumption [17,18]. For the distant future, a practical economist engaged in the environmental debate might try to use, as the forward discount rate, the average of historical interest rates which occurred in the past 200 years (2.7% in stable countries [18]) or take the average of Wall Street forward looking models that price bonds of maturity as long as 30 years. However, we have recently shown [18] that, due to historical fluctuations of short real interest rates, the appropriate rate is considerably below these averages.

Moreover, the presence of fluctuations can dramatically alter the functional form of the discount function. If interest rates follow a geometric random walk, for example, Farmer and Geanakoplos [19] (see also Ref. [20]) have shown that the

*Doyne.Farmer@inet.ox.ac.uk
†john.geanakoplos@yale.edu
‡jaume.masoliver@ub.edu
§miquel.montero@ub.edu
‖josep.perello@ub.edu

discount function decays as a power law of the form $t^{-1/2}$. In contrast to the exponential function, this is not integrable on $(0,\infty)$, underscoring how important the effect of fluctuations can be.

One purpose of this paper is to present a comprehensive approach to the problem of discounting and the modeling of interest rates within the frames of reference of statistical mechanics and econophysics. Since we have very recently treated discounting for environmental economics problems in which real rates have a non-negligible probability of taking negative values [18], a second objective consists in extending our former model to embrace other situations such as those appearing in purely financial settings where bond rates never become negative.

To understand how discounting depends on the random process used to characterize interest rates, we study three different models which appear ubiquitously in the literature [14–16]. The models are based on the Ornstein-Uhlenbeck (OU) process [11,21], the Feller process [12,22], and the log-normal process [13,23]. In two of them (Feller and log-normal) rates cannot take negative values while in the OU process they can be either positive or negative.

Which model is most appropriate depends on the problem under study and its context. Thus if we deal with environmental problems we should use real interest rates which are nominal interest rates corrected by inflation [17,18]. In Ref. [18] we have done a thorough empirical study of real interest rates across 14 countries covering 87 to 318 years and the data clearly show that in many epochs and for all countries real interest rates frequently become negative, often by substantial amounts and for long periods of time. In environmental problems we are, therefore, led (by tractability as well) to the OU model. Nevertheless, financial settings generally use nominal interest rates which are positive and, therefore, either the Feller or the log-normal processes are more appropriate, as has long been recognized [14–16].

The paper is organized as follows. In Sec. II we present a general analysis of the problem of discounting aimed at the nonspecialist and within an approach resembling that of physics. In Secs. III–V we obtain exact expressions of the discount functions and their asymptotic approximations for three relevant models of interest rates. Concluding remarks are presented in Sec. VI, whereas more technical details are in appendices.

II. GENERAL ANALYSIS

In economics the increment at a given time of the quantity of wealth, exemplified by some magnitude $M(t)$, is assumed to depend linearly on the quantity itself and the duration of the variation. For a continuous and instantaneous variation one writes:

$$dM(t) \propto M(t)dt. \tag{2.1}$$

This is a phenomenological law based not only on the empirical fact that the bigger $M(t)$, the greater its variation at a given time, but also on the simplifying assumption that the increment is linear in $M(t)$ and not, for instance, quadratic. [Note that linearity is equivalent to assuming that the interest rate is independent of $M(t)$]. The growth law (2.1) has nonetheless been justified by Vasicek [11] on purely financial grounds within the framework of a general theory of the term structure of interest rates (see also Ref. [24]).

In the simplest situation Eq. (2.1) represents a completely linear law with direct proportionality

$$dM(t) = rM(t)dt, \tag{2.2}$$

where r is the rate and is measured in units of $1/(\text{time})$. Now the growth law is readily integrated, giving

$$M(t) = e^{r(t-t_0)}M(t_0), \tag{2.3}$$

which yields an exponential growth connecting wealth at some initial time t_0, that is to say, the present time (which, in our case and without loss of generality, can be taken equal to zero) and wealth at any future time t.

Before proceeding further we recall that the growth law (2.1), often in the simplest version (2.2), appears in numerous branches of physical and social sciences. Thus, for example, in radioactivity, if $N(t)$ is the number of active nuclei at time t, then the usual hypothesis is that this number decreases as

$$dN(t) = -\lambda N(t)dt,$$

where $\lambda > 0$ is the decay constant. Similar considerations apply to other situations, as they are found in chemical reactions, population dynamics, as in many other places.

In economics, discounting refers to the process of connecting wealth at different times. Specifically the discount function, which we denote by $d(t)$, is defined by

$$d(t) \equiv \frac{M(0)}{M(t)}, \tag{2.4}$$

so $M(0) = d(t)M(t)$ in accordance with the fact that discounting specifically refers to weighting the future (at some time $t > 0$) relative to the present (at $t = 0$).

In the simplest case of Eq. (2.3) the discount function is given by the decreasing exponential

$$d(t) = e^{-rt}, \tag{2.5}$$

where $r > 0$ is the interest rate. However, this simple form of discount, in which the interest rate is always constant, is unrealistic. A first generalization consists in assuming rates to be known functions of time $r(t)$. In such a case the growth law (2.2) is replaced by

$$dM(t) = r(t)M(t)dt$$

and discount is given by

$$d(t) = \exp\left[-\int_0^t r(t')dt'\right]. \tag{2.6}$$

Obviously, if $r(t) = r$ is constant we recover the simple exponential decay of Eq. (2.5).

However, the assumption of rates being given by constants or by deterministic functions of time is unreasonable, at least over long periods of time. Financial interest rates are typically described as random, as the many models for stochastic interest rates appearing in the literature show [14–16]. Population dynamics are subject to random influences, as are chemical reactions and other physical processes where rates appear.

We therefore assume that $r(t)$ is a random function of time. This naturally means that discounting is also random. In these circumstances the effective discount function is defined as the average of $d(t)$,

$$D(t) = \left\langle \exp\left[-\int_0^t r(t')dt'\right]\right\rangle, \tag{2.7}$$

taken over all possible realization of $r(t)$ [11]. The function $r(t)$ can, in principle, be any random variable. However, the most common assumption is that rates are diffusion processes [14–16]. Indeed, a natural simplifying assumption is that $r(t)$ is a Markovian processes with continuous paths, that is, a diffusion processes [11]. Therefore, rates are solutions to stochastic differential equations of the form

$$dr(t) = f[r(t)]dt + g[r(t)]dW(t), \tag{2.8}$$

where $f(r)$ and $g(r)$ do not depend explicitly on time, $W(t)$ is the Wiener process, and the stochastic differential equation is interpreted in the Itô sense.

Note that we assume that drift $f(r)$ and noise intensity $g(r)$ do not depend explicitly on time [the time dependence is only implicit through $r = r(t)$]. This means that the interest rate process is time homogeneous and stationary. This constitutes an idealization because real markets do not seem to be stationary, at least over long periods of time [18]. In future work we will try to relax this restriction.

Defining the auxiliary random process

$$x(t) = \int_0^t r(t')dt', \tag{2.9}$$

the equivalent discount function can be written as

$$D(t) = \langle e^{-x(t)}\rangle.$$

Therefore,

$$D(t) = \int_{-\infty}^{\infty} dr \int_{-\infty}^{\infty} e^{-x} p(x,r,t|r_0)dx, \tag{2.10}$$

where $p(x,r,t|r_0)$ is the probability density function (PDF) of the bidimensional diffusion process $[x(t),r(t)]$.

From Eqs. (2.8) and (2.9) we see that $[x(t),r(t)]$ is defined by the following pair of stochastic differential equations:

$$\begin{aligned} dx &= rdt, \\ dr &= f(r)dt + g(r)dW(t). \end{aligned} \tag{2.11}$$

Therefore, the joint density obeys the Fokker-Planck equation (FPE) [25]

$$\frac{\partial p}{\partial t} = -r\frac{\partial p}{\partial x} - \frac{\partial}{\partial r}[f(r)p] + \frac{1}{2}\frac{\partial^2}{\partial r^2}[g^2(r)p], \tag{2.12}$$

with the initial condition

$$p(x,r,0|r_0) = \delta(x)\delta(r - r_0). \tag{2.13}$$

It turns out to be most convenient for further developments to work with the characteristic function instead of the PDF, the former defined as the Fourier transform of the latter:

$$\tilde{p}(\omega_1,\omega_2,t|r_0) = \int_{-\infty}^{\infty} e^{-i\omega_1 x}dx \times \int_{-\infty}^{\infty} e^{-i\omega_2 r} p(x,r,t|r_0)dr. \tag{2.14}$$

One of the chief advantages of working with the characteristic function $\tilde{p}$ is that obtaining the effective discount is straightforward. Indeed, comparison of Eqs. (2.10) and (2.14) shows that

$$D(t) = \tilde{p}(\omega_1 = -i,\omega_2 = 0,t). \tag{2.15}$$

To our knowledge this formalism based on Fourier analysis is not used in the economic literature of discounting, where the usual approach consists in applying the Feynman-Kac formula [26] or in the use of Fokker-Planck methods [27,28] and even quantum-field-theoretical methods [29].

In the following sections we try to understand how discounting depends on the random process used to model interest rates. The models chosen are the OU process [21], the Feller process [22], and the log-normal process [23]. All of them are widely used in the interest rate literature [11–16]. In the OU model rates can be either positive or negative while in Feller and log-normal models rates cannot take on negative values.

III. THE ORNSTEIN-UHLENBECK MODEL

In the theory of financial interest rates the OU model was proposed in Ref. [11] and it is sometimes referred to as the Vasicek model. The model is a diffusion process characterized by linear drift and constant noise intensity

$$dr(t) = -\alpha[r(t) - m]dt + kdW(t), \tag{3.1}$$

where $r(t)$ is the real interest rate, $W(t)$ is the Wiener process. The parameter m, sometimes referred to as "normal level," is a mean value to which the process reverts, $k > 0$ is the amplitude of fluctuations, and $\alpha > 0$ is the strength of the reversion to the mean. These parameters must be estimated from empirical data [17,18].

In Appendix A we show that in the stationary regime the solution to Eq. (3.1) is

$$r(t) = m + k\int_{-\infty}^{t} e^{-\alpha(t-t')}dW(t'), \tag{3.2}$$

which, due to the unbounded variation of the Wiener process, clearly shows that rates can be negative. In fact, we have proved that the probability for $r(t)$ to be negative is [18]

$$P^{(-)} = \tfrac{1}{2}\mathrm{Erfc}(m\sqrt{\alpha/k^2}).$$

Moreover, it follows from Eq. (3.2) that the normal level is the (stationary) mean value of the process:

$$\langle r(t)\rangle = m,$$

where the average must be taken in the stationary regime. We also show in Appendix A that the (stationary) correlation function, $C(\tau) = \langle[r(t+\tau) - m][r(t) - m]\rangle$, of the OU process is

$$C_{\mathrm{OU}}(\tau) = \left(\frac{k^2}{2\alpha}\right)e^{-\alpha\tau}, \tag{3.3}$$

showing that α^{-1} is the autocorrelation time and

$$\sigma_{\mathrm{OU}}^2 = \frac{k^2}{2\alpha} \tag{3.4}$$

is the volatility.

Let us next focus on the joint PDF $p(x,r,t|r_0)$. Since now $f(r) = -\alpha(r-m)$ and $g(r) = k$ the FPE (2.12) reads

$$\frac{\partial p}{\partial t} = -r\frac{\partial p}{\partial x} + \alpha\frac{\partial}{\partial r}[(r-m)p] + \frac{1}{2}k^2\frac{\partial^2 p}{\partial r^2}, \tag{3.5}$$

with the initial condition given by [cf. Eq. (2.13)]

$$p(x,r,0|r_0) = \delta(x)\delta(r-r_0). \tag{3.6}$$

Fourier transforming Eqs. (3.5) and (3.6) results in a simpler problem for the characteristic function

$$\frac{\partial \tilde{p}}{\partial t} = (\omega_1 - \alpha\omega_2)\frac{\partial \tilde{p}}{\partial \omega_2} - \left(i\alpha m\omega_2 + \frac{k^2}{2}\omega_2^2\right)\tilde{p}, \tag{3.7}$$

with

$$\tilde{p}(\omega_1,\omega_2,0|r_0) = e^{-i\omega_2 r_0}. \tag{3.8}$$

In Appendix A we prove that the exact solution of this initial-value problem is given by a Gaussian density from which we can readily obtain the effective discount function $D(t)$ after setting $\omega_1 = -i$ and $\omega_2 = 0$ in that Gaussian distribution [cf. Eq. (2.15)]. The final result reads (see, e.g., Ref. [14])

$$\begin{aligned}\ln D(t) = &-\frac{r_0}{\alpha}(1-e^{-\alpha t}) \\ &+ \frac{k^2}{2\alpha^3}\left[\alpha t - 2(1-e^{-\alpha t}) + \frac{1}{2}(1-e^{-2\alpha t})\right] \\ &- m\left[t - \frac{1}{\alpha}(1-e^{-\alpha t})\right].\end{aligned} \tag{3.9}$$

The exponential terms in Eq. (3.9) are only significant for times smaller than the autocorrelation time of the rate, $t < \alpha^{-1}$. At longer times the exponential terms are negligible. Moreover, as time increases even constant terms are negligible. We thus have

$$D(t) \simeq e^{-\varrho t}, \tag{3.10}$$

where

$$\varrho = m - k^2/2\alpha^2, \tag{3.11}$$

is the long-run discount rate.

Notice that ϱ is always smaller than the average interest rate m by an amount that depends on the square of the product, $k\alpha^{-1}$, of the noise intensity by the correlation time of the process, the latter indicating the persistence of fluctuations. Thus, for a given k, a long persistence (i.e., α small) diminishes considerably the long-run discount rate ϱ with respect to the mean value m. The same occurs with a fixed correlation time but sufficiently intense fluctuations (i.e., k large). In other words, by varying k and α, the long-run rate can take on any value less than m, including negative values, while at the same time the volatility—which depends on $k^2\alpha^{-1}$, Eq. (3.4)—can also be made to take any arbitrary positive number. That is, by choosing the appropriate (k,α), we can make ϱ arbitrarily far below m, with σ arbitrarily small.

On the other hand, we also see that the long-run interest rate may be negative. How is it possible for ϱ to be negative and thus for the discount function $D(t)$ to increase? This is easy to understand when there are persistent periods of negative real interest rates $r(t)$. Computation of the discount function $D(t)$ in Eq. (2.7) involves an average over exponentials rather than the exponential of an average. As a result, periods where interest rates are negative are greatly amplified and can easily dominate periods where interest rates are large and positive, even if the negative rates are rarer and weaker. It does not take many such periods to produce long-run exponential growth of $D(t)$.

To summarize, the long-run discount rate can be much lower than the mean and indeed can correspond to low interest rates that are rarely observed. For environmental problems this dramatically illustrates the imprudence of assuming that the average real interest rate is the correct long-run discount rate.

IV. THE FELLER MODEL

In this section and the next we present two models in which rates cannot be negative.

One of the most accepted models for interest rates is the so-called Cox-Ingersoll-Ross (CIR) model [12] in which rates are assumed to follow the Feller process [22] defined by the stochastic differential equation [12,22,30]

$$dr(t) = -\alpha[r(t)-m]dt + k\sqrt{r(t)}dW(t). \tag{4.1}$$

The process represents a diffusion model with linear drift $f(r) = -\alpha(k-m)$ and state-dependent noise intensity $g(r) = k\sqrt{r}$. The quantities α, m, and k are positive parameters. All differentials are interpreted in the sense of Ito. Note that since the diffusion coefficient in one-dimensional diffusion processes is given by the square of the noise intensity, we see that the Feller process has a linear diffusion vanishing at the origin. This turns the origin into a singular boundary.

As in the OU model, the linear drift results in a restoring force which, in the absence of noise, makes the process decay toward the "normal level" m. On the other hand, the state-dependent noise intensity $k\sqrt{r}$ for large values of r magnifies the effect of noise while when r goes to zero this effect vanishes. Therefore, as the process approaches the origin the drift drags r towards m. Hence, since $m > 0$, starting at some positive value $r_0 > 0$ the process cannot attain negative values, with the overall result that the Feller process remains always positive.

Before proceeding further we will briefly review the main characteristics of the process and the singular character of the origin. Like the OU process, the Feller process has a stationary limit and the reverting level m is the average value in the stationary regime. The correlation function is also given by an exponential decreasing in time as in Eq. (3.3) (see Appendix B)

$$C_F(\tau) = \left(\frac{mk^2}{2\alpha}\right)e^{-\alpha\tau}. \tag{4.2}$$

Note that α^{-1} is again the correlation time but, contrary to the OU process, the volatility

$$\sigma_F^2 = \frac{mk^2}{2\alpha} \tag{4.3}$$

depends on the normal level as well.

The Feller process is not Gaussian and its PDF is given by [22,30]

$$p(r,t|r_0) = \frac{2\alpha/k^2}{1-e^{-\alpha t}}\left(\frac{re^{\alpha t}}{r_0}\right)^{(\theta-1)/2} \times \exp\left\{-\left(\frac{2\alpha}{k^2}\right)\frac{r+r_0e^{-\alpha t}}{1-e^{\alpha t}}\right\} \times I_{\theta-1}\left[\left(\frac{4\alpha}{k^2}\right)\frac{\sqrt{rr_0e^{-\alpha t}}}{1-e^{-\alpha t}}\right], \quad (4.4)$$

where $I_{\theta-1}[\cdot]$ is the modified Bessel function and

$$\theta = \frac{2\alpha m}{k^2} \quad (4.5)$$

is a positive and dimensionless constant which combines all the parameters of the model into a single expression.

The behavior of the process at the singular boundary $r = 0$, which we have discussed above in a qualitative manner, can be seen in a more quantitative way from the analysis of the PDF near the origin. Thus by expanding Eq. (4.4) around $r = 0$ it is not difficult to show [30] that

$$p(0,t|x_0) = \begin{cases} \infty & \theta < 1 \\ 0 & \theta > 1, \end{cases} \quad (4.6)$$

for all t, with a time-dependent finite value for $p(0,t|r_0)$ when $\theta = 1$.

This proves that when $\theta > 1$ the probability for the Feller process to reach the value $r = 0$ is zero, whereas if $\theta \leqslant 1$ this probability is greater than zero. In other words, if $\theta \leqslant 1$, then the origin is an accessible boundary but if $\theta > 1$, then it is not. We, therefore, conclude that in the CIR model rates never vanish (in probability) as long as

$$\frac{2\alpha m}{k^2} > 1.$$

The Feller process is stationary and its stationary PDF is given by the Gamma distribution which is readily obtained after taking the limit $t \to \infty$ in Eq. (4.4):

$$p_{\rm st}(r) = \frac{(2\alpha/k^2)^\theta}{\Gamma(\theta)} r^{\theta-1} e^{-(2\alpha/k^2)r}. \quad (4.7)$$

In the Feller model, the joint density of the discounting process $(x(t),r(t))$ defined in Eq. (2.11) obeys the Fokker-Planck equation (FPE) [cf. Eqs. (2.12) and (2.13)],

$$\frac{\partial p}{\partial t} = -r\frac{\partial p}{\partial x} + \alpha\frac{\partial}{\partial r}[(r-m)p] + \frac{k^2}{2}\frac{\partial^2}{\partial r^2}(rp), \quad (4.8)$$

with the initial condition

$$p(x,r,0|r_0) = \delta(x)\delta(r-r_0). \quad (4.9)$$

The joint Fourier transform, Eq. (2.14), turns Eqs. (4.8) and (4.9) into a more manageable problem:

$$\frac{\partial \tilde{p}}{\partial t} = \left(\omega_1 - \alpha\omega_2 - i\frac{k^2}{2}\omega_2^2\right)\frac{\partial \tilde{p}}{\partial \omega_2} - i\alpha m\omega_2\tilde{p}, \quad (4.10)$$

$$\tilde{p}(\omega_1,\omega_2,0|r_0) = e^{-i\omega_2 r_0}. \quad (4.11)$$

Equation (4.10) is a linear partial differential equation of first order whose solution can be obtained by the method of characteristics [31]. This is done in Appendix B.

As shown in Eq. (2.15), once we know the form of the joint characteristic function, $\tilde{p}(\omega_1,\omega_2,t|r_0)$, the discount function is readily obtained by setting $\omega_1 = -i$ and $\omega_2 = 0$. In Appendix B we also show that (see, e.g., [14])

$$D(t) = \left[\frac{2\lambda e^{-(\lambda-\alpha)t/2}}{(\lambda+\alpha)+(\lambda-\alpha)e^{-\lambda t}}\right]^\theta \times \exp\left\{-\frac{2(1-e^{-\lambda t})r_0}{(\lambda+\alpha)+(\lambda-\alpha)e^{-\lambda t}}\right\}, \quad (4.12)$$

where θ is defined in Eq. (4.5) and

$$\lambda = \sqrt{\alpha^2 + 2k^2}. \quad (4.13)$$

Notice that $\lambda > \alpha$ and the time scale represented by λ^{-1} is smaller than the correlation time α^{-1}.

As time increases (in fact, when $\lambda t \gg 1$) the effective discount (4.12) reduces to

$$D(t) \simeq e^{-\varrho t} \quad (4.14)$$

$(t \to \infty)$, where

$$\varrho = \tfrac{1}{2}(\lambda-\alpha)\theta \quad (4.15)$$

is the long-run discount rate of the Feller model. Substituting for Eqs. (4.5) and (4.13) this can be written as

$$\varrho = \frac{2m}{1+\sqrt{1+2k^2/\alpha^2}}, \quad (4.16)$$

which clearly shows that

$$\varrho < m.$$

Therefore, like the OU model, the Feller long-run discount rate is smaller than the stationary average rate by an amount that also depends on the square of the product $k\alpha^{-1}$. Notice that, again, either a long persistence (α small) or an increase of the noise intensity (k large) diminish the long-run discount rate.

V. THE LOG-NORMAL MODEL

In economics the log-normal process, also called the geometric Brownian motion, has been widely used for the modeling of price dynamics [23]. It can be written as

$$\frac{dr}{r} = \alpha dt + k dW(t), \quad (5.1)$$

where r is the interest rate, α and k are constant parameters, and α may be positive or negative, whereas k is always positive and $W(t)$ is the Wiener process. Equation (5.1) can be integrated at once, yielding

$$r(t) = r_0 \exp\left\{\left(\alpha - \frac{k^2}{2}\right)t + kW(t)\right\}, \quad (5.2)$$

showing that $r(t)$ is never negative ($r_0 > 0$). Therefore, the log-normal model is more suited for modeling nominal interest rates in finance, which are never negative, than for modeling real rates in environmental economics. Contrary to OU and Feller models, the log-normal process does not show reversion

to the mean. Indeed, as t increases we see from Eq. (5.2) that the rate either diverges when $\alpha > 0$ or goes to zero if $\alpha < 0$. In an equivalent way one can also show from Eq. (5.2) that the mean and variance of the process are [14]

$$\langle r(t)\rangle = r_0 e^{\alpha t}, \quad \mathrm{Var}[r(t)] = r_0^2 e^{2\alpha t}(e^{k^2 t} - 1).$$

The discount associated with the log-normal process model was studied in 1978 [13] and in finance is usually known as the Dothan model. Because it allows for analytical treatment it is one of the models used in the literature [14] and we will compute the discounting function and discuss some of its interesting asymptotic properties which, to our knowledge, have not been described before (see, however, Refs. [19,20]).

For the log-normal model it is possible to obtain an exact expression of the discount function for its (time) Laplace transform:

$$\hat{D}(\sigma) = \int_0^\infty e^{-\sigma t} D(t) dt.$$

Following lengthy calculations summarized in Appendix C we obtain

$$\hat{D}(\sigma) = \frac{2\Gamma(\beta)}{k^2\Gamma(2\beta+\gamma)} \int_0^\infty \frac{e^{-(2r_0\zeta/k^2 + 1/\zeta)}}{\zeta^{\beta+\gamma}} \times F(\beta, 2\beta+\gamma, \zeta^{-1}) d\zeta, \tag{5.3}$$

where $F(a,b,x)$ is a Kummer function [32],

$$\gamma = 2(1 - \alpha/k^2) \tag{5.4}$$

and

$$\beta = \tfrac{1}{2}[1 - \gamma + \sqrt{(1-\gamma)^2 + 4\sigma}]. \tag{5.5}$$

The expression in Eq. (5.3) is as far as we can go from an analytical point of view because the exact inversion of Eq. (5.3) to get $D(t)$ is difficult and unpractical because the exact $D(t)$ is clumsy and not amenable even to numerical work or asymptotic approximations [13,14]. On the other hand, $\hat{D}(\sigma)$ allows us to get, in a relatively simple way, asymptotic expressions of the discount function valid for large times, which is of interest for environmental problems. Another way of getting the complete behavior of $D(t)$ for all t would be to perform the Laplace inversion of Eq. (5.3) numerically but this would need the setting of parameters α and k to fixed values with subsequent loss of generality and scope.

Long-time asymptotic expressions of $D(t)$ are easily derived using Tauberian theorems which relate the small σ behavior of $\hat{D}(\sigma)$ with the long-time behavior of $D(t)$ [33]. In Appendix C we show that as $t \to \infty$

$$D(t) \sim \begin{cases} \text{const} & \alpha < k^2/2, \\ e^{-\varrho t} & \alpha > k^2/2, \\ t^{-1/2} & \alpha = k^2/2. \end{cases} \tag{5.6}$$

The asymptotic form of the discount function thus depends on the values taken by the ratio α/k^2—the "signal-to-noise" ratio of the problem—between the strength of the deterministic drift α and the amplitude of fluctuations represented by $k^2/2$.

(i) The case $k^2/2 > \alpha$ corresponds to strong fluctuations, where the noise intensity $k^2/2$ is greater than the drift parameter α ruling deterministic motion (i.e., the signal is weaker than the noise). In this case the discount asymptotes to a constant value whose value is given in Eq. (C23) of Appendix C.

(ii) The case $k^2/2 < \alpha$ corresponds to milder fluctuations for which the signal is stronger than noise. In such a case the discount function has the expected exponential decay

$$D(t) \sim e^{-\varrho t}, \tag{5.7}$$

with a long-run rate of discount given by [cf. Eq. (C31) of Appendix C]

$$\varrho = \frac{1}{\delta}\left(\alpha - \frac{k^2}{2}\right), \tag{5.8}$$

where $0 < \delta < 1$ (see Fig. 1) is a positive numerical factor which only depends on the ratio $2\alpha/k^2$ and reads [cf. Eq. (C36) of Appendix C]

$$\delta = \psi(2\alpha/k^2) + \frac{1}{2\alpha/k^2 - 1}, \tag{5.9}$$

where $\psi(\cdot)$ is the digamma function.

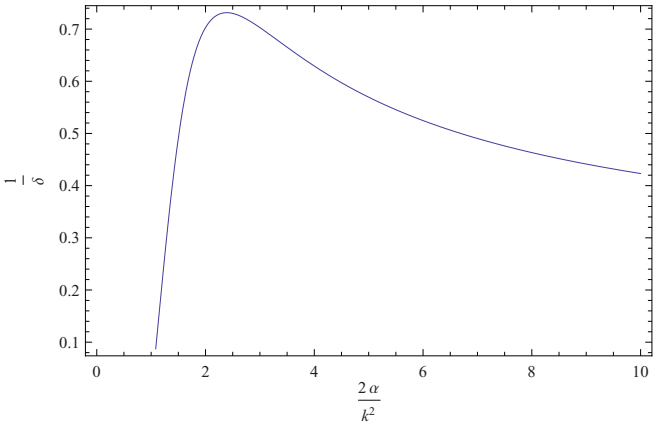

FIG. 1. (Color online) The factor $1/\delta$, Eq. (5.9), as a function of the signal-to-noise ratio $2\alpha/k^2 > 1$. The factor attains the approximate maximum value 0.73 for $2\alpha/k^2 \simeq 2.4$.

Let us write Eq. (5.7) in a more suggestive form. From Eq. (5.2) we see that

$$\left\langle \ln \frac{r(t)}{r_0} \right\rangle = \left(\alpha - \frac{k^2}{2}\right) t.$$

Hence,

$$D(t) \sim \exp\left[-\frac{1}{\delta}\left\langle \ln \frac{r(t)}{r_0}\right\rangle\right] \tag{5.10}$$

($t \to \infty$ and $k^2/2 < \alpha$).

Note that the average $\langle \ln r(t)/r_0\rangle$ is what a practitioner would take as an estimate of the discount rate up to time t within the log-normal model. The analytical result (5.10) shows that the actual long-run rate of the model is a fraction of the average rate. Indeed, in Fig. 1 we plot the factor $1/\delta$ in terms of the signal-to-noise ratio $2\alpha/k^2$ [cf. Eq. (5.9)] and we find the maximum value attained by that factor to be around $1/\delta \approx 0.73$ at $2\alpha/k^2 \approx 2.4$. In other words the long-run discount rate is at most 73% of the average rate. In this way when $2\alpha/k^2 > 1$ the log-normal model follows a similar pattern to that of the OU and Feller models: In all of them the long-run rate is smaller than the average rate.

(iii) The critical case $\alpha = k^2/2$ leads to the hyperbolic discount function as obtained by Farmer and Geanakoplos [19,20].

VI. CONCLUDING REMARKS

Financial economists have developed a large number of models of interest rate processes to enable them to price bonds and other cash flows. In these models interest rates are described by positive random processes since financial interest rates never (or very rarely) go negative. Although the models could in principle be extended to arbitrary horizons, they have only been studied carefully over time horizons of up to 30 years, since bonds are rarely issued for periods longer than this.

On the other hand, environmental economists are interested in the real behavior of the economic growth over much larger horizons, in contrast to financial economists, who are typically more interested in nominal rates over shorter horizons. Their behavior essentially differs due to the fact that real rates can take on negative values. Taking nominal rates corrected by inflation as a proxy of economic growth, with a thorough empirical study on many countries [18], we demonstrated that real interest rates are negative around 20% of the time (including at the time of writing this paper).

To understand how discounting depends on the random process used to characterize interest rates we have studied three different models and obtained exact analytical expressions for the discount function. The three models describe to varying degree a number of relevant characteristics observed in rates, while being simple enough to allow for complete analytical treatment.

The first model is the Ornstein-Uhlenbeck process (Vasicek model) which allows for negative rates and is therefore suitable for pricing environmental problems. The model has a stationary probability distribution and exhibits reversion to the mean, which means that the process tends to return to its average stationary value.

The second and third models considered are given by the Feller and log-normal processes, respectively. For these processes rates cannot be negative. The Feller process (CIR model) has reversion to the mean and a stationary probability distribution. It is one of the most popular models in finance. On the other hand, the log-normal process (Dotham model) does not have reversion to the mean and does not have a stationary distribution. Despite these shortcomings the process has also been used in the financial literature mainly because it is positive and allows for analytical treatment [14].

Using Fourier methods we have been able to reproduce the exact discount functions for all these models, which one otherwise finds scattered through a vast literature and derived by other methods [34]. One of our most interesting results is in obtaining the long-run discount rate for each model. We have thus proved that asymptotically the discount function behaves, for all models, as a decaying exponential

$$D(t) \sim e^{-\varrho t}$$

($t \to \infty$), where ϱ is the long-run discount rate which, according to each model, is given by

$$\varrho = \begin{cases} m - k^2/2\alpha^2 & \text{(OU)} \\ 2m/[1+\sqrt{1+2k^2/\alpha^2}] & \text{(Feller)} \end{cases}$$

and

$$\varrho t = \frac{1}{\delta}(\alpha - k^2/2)t = \frac{1}{\delta}\langle \ln r(t)/r_0 \rangle \quad \text{(log-normal)}.$$

In these expressions m is the mean value (in OU and Feller models), α is the strength of the drift (for the OU and Feller processes α^{-1} also represents the correlation time), and k is the intensity of the noise. The expression for ϱ appearing in the log-normal case is only valid when $2\alpha/k^2 > 1$, otherwise the discount function has a different functional form [see Eq. (5.6)] and the number $\delta > 1$ is [cf. Eq. (5.9) and Fig. 1]

$$\delta = \psi(2\alpha/k^2) + \frac{1}{2\alpha/k^2 - 1}.$$

A remarkable property is that the long-run discount rate is always less than the average of past interest rates by an amount that depends on α and k. This general statement is indeed a direct consequence of Jensen's inequality, which states that the average of a convex function is greater than or equal to the function of the average, i.e., $\langle f(X)\rangle \geqslant f(\langle X\rangle)$. Assuming f to be the decreasing exponential and X the integral process $x(t)$ defined in Eq. (2.9), it follows immediately that the long-run rate ϱ must be always less than or equal to the average rate. However, we have gone further than this qualitative statement because we have quantified the difference. Thus, in the OU and Feller processes the (stationary) average interest rate is m and, in both cases, the long-run rate ϱ attains the maximum value (which equals m) when $k/\alpha \to 0$ and then decreases monotonically as k/α increases (into negative ϱ's in the OU model and towards $\rho = 0$ in Feller's). The case of the log-normal model slightly differs, for the process is not stationary and does not attain a stationary mean value. In any case, the long-run discount for the log-normal model is always smaller than the average of the logarithm of interest rates and the maximum discount rate is around 73% of the average value (cf. Fig. 1). All of this confirms the Weitzman-Gollier principle [35,36] that the volatility and persistence of interest rates lower long-run discounting.

As mentioned in Sec. I we have observed significant differences between long-run and average rates in historical data across 14 countries covering 87 to 318 years [18]. Thus, for instance, in the United Kingdom the historical average rate over more than 300 years is 3.3% while the long-run discount rate is 2.8% (where the latter figure has been obtained using the OU model with all parameters estimated from data; see Ref. [18] for more details and countries). The long-run rate of 2.8% is very close to that recently obtained by Giglio *et al.* [37]. Using data of housing markets in United Kingdom and Singapore they estimate an annual discount rate of 2.6%. Finally, the average historical rate of stable countries is 2.7%, whereas, due to fluctuations, the long-run discounting rate is 2.1%, which is around 22% smaller than the historical average [18].

We conclude on purely theoretical grounds that those who wish to do cost benefit analysis, especially over long periods of time as in environmental economics (but in finance as well), must not hastily use a discount rate set equal to the historical average of the short-run rate. Doing so biases

discount calculations in favor of the present and against interventions that may protect the future.

ACKNOWLEDGMENTS

J.M., M.M., and J.P. acknowledge partial from MINECO under Contracts No. FIS2013-47532-C3-2-P and No. FIS2009-09689 and also from Generalitat de Catalunya under Contract No. 2014 SGR 608.

APPENDIX A: THE ORNSTEIN-UHLENBECK PROCESS

The solution to Eq. (3.1) is

$$r(t) = r(t_0)e^{-\alpha(t-t_0)} + m[1-e^{-\alpha(t-t_0)}] + k\int_{t_0}^{t} e^{-\alpha(t-t')}dW(t'),$$

where t_0 is an arbitrary initial time. In the stationary regime the process is supposed to have begun in the infinite past so $t_0 \to -\infty$ and

$$r(t) = m + k\int_{-\infty}^{t} e^{-\alpha(t-t')}dW(t'). \tag{A1}$$

Since $\langle dW(t')\rangle = 0$ we therefore have

$$\langle r(t)\rangle = m$$

and the normal level m is the stationary mean value.

The correlation function is defined by

$$C(\tau) = \langle [r(t+\tau) - \langle r(t)\rangle]\,[r(t) - \langle r(t)\rangle]\rangle\,.$$

In the stationary regime $r(t)$ is given by Eq. (A1) and $\langle r(t)\rangle = m$. Hence

$$C_{\rm OU}(\tau) = k^2 e^{-\alpha(2t+\tau)}\int_{-\infty}^{t+\tau} e^{\alpha t_1}\int_{-\infty}^{t} e^{\alpha t_2}\langle dW(t_1)dW(t_2)\rangle.$$

Since $\langle dW(t_1)dW(t_2)\rangle = \delta(t_1-t_2)dt_1dt_2$ we finally obtain

$$C_{\rm OU}(\tau) = \left(\frac{k^2}{2\alpha}\right)e^{-\alpha\tau},$$

which is Eq. (3.3).

1. The joint density

Let us prove that the solution to the initial-value problem (3.7) and (3.8) is given by the Gaussian function

$$\tilde p(\omega_1,\omega_2,t) = \exp\left\{-A(\omega_1,t)\omega_2^2 - B(\omega_1,t)\omega_2 - C(\omega_1,t)\right\}. \tag{A2}$$

In effect, substituting (A2) into (3.7), identifying like powers in ω_2, and taking into account the initial condition (3.8), we find that the unknown functions A, B, and C satisfy the following set of ordinary differential equations:

$$\dot A = -2\alpha A - k^2/2, \quad A(\omega_1,0) = 0,$$
$$\dot B = -\alpha B + 2\omega_1 A - i\alpha m, \quad B(\omega_1,0) = ir_0,$$

and

$$\dot C = \omega_1 B, \quad C(\omega_1,0) = 0.$$

These equations are sequentially integrated, giving

$$A(\omega_1,t) = \frac{k^2}{4\alpha}(1-e^{-2\alpha t}),$$
$$B(\omega_1,t) = ir_0e^{-\alpha t} + \frac{k^2\omega_1}{2\alpha^2}(1-2e^{-\alpha t}+e^{-2\alpha t}) + im(1-e^{-\alpha t}),$$

and

$$C(\omega_1,t) = i\omega_1 r_0\frac{1}{\alpha}(1-e^{-\alpha t}) + \frac{k^2\omega_1^2}{2\alpha^3}\left[\alpha t - 2(1-e^{-\alpha t}) + \frac{1}{2}(1-e^{-2\alpha t})\right] + im\omega_1\left[t - \frac{1}{\alpha}\left(1-e^{-\alpha t}\right)\right]. \tag{A3}$$

From Eq. (2.15) we see that the effective discount is given by the characteristic function, $\tilde p(\omega_1,\omega_2,t|r_0)$, evaluated at the points $\omega_1 = -i$ and $\omega_2 = 0$. Thus using Eqs. (A2) and (A3) we finally get Eq. (3.9).

APPENDIX B: THE FELLER PROCESS

We write Eq. (4.1) in the form

$$\dot r = -\alpha(r-m) + k\sqrt{r}\xi(t), \tag{B1}$$

where $\xi(t) = dW(t)/dt$ is Gaussian white noise such that

$$\langle\xi(t)\rangle = 0, \quad \langle\xi(t)\xi(t')\rangle = \delta(t-t'). \tag{B2}$$

In the stationary regime (i.e., $t_0 \to -\infty$) the formal solution of Eq. (B1) can be written as

$$r(t) = m + k\int_{-\infty}^{t} e^{-\alpha(t-t')}\sqrt{r(t')}\xi(t')dt', \tag{B3}$$

from which we see that the (stationary) mean value of $r(t)$ is

$$\langle r(t)\rangle = m, \tag{B4}$$

where we have taken into account that $r(t)$ and $\xi(t)$ are uncorrelated (Itô convention), so

$$\langle\sqrt{r(t')}\xi(t')\rangle = \langle\sqrt{r(t')}\rangle\langle\xi(t')\rangle = 0.$$

We can also evaluate from Eq. (B3) the (stationary) correlation function defined as

$$C(\tau) = \langle [r(t+\tau)-m][r(t)-m]\rangle.$$

Bearing in mind that $r(t)$ and $\xi(t)$ are uncorrelated, from Eqs. (B2)–(B4) we have

$$\begin{aligned} C_F(\tau) &= k^2\int_{-\infty}^{t+\tau} dt_1 e^{-\alpha(t+\tau-t_1)}\int_{-\infty}^{t} dt_2 e^{-\alpha(t-t_2)} \\ &\quad\times\langle\sqrt{r(t_1)}\sqrt{r(t_2)}\xi(t_1)\xi(t_2)\rangle \\ &= k^2\int_{-\infty}^{t+\tau} dt_1 e^{-\alpha(t+\tau-t_1)}\int_{-\infty}^{t} dt_2 e^{-\alpha(t-t_2)} \\ &\quad\times\langle\sqrt{r(t_1)}\sqrt{r(t_2)}\rangle\delta(t_1-t_2) \\ &= k^2 m e^{-\alpha(2t+\tau)}\int_{-\infty}^{t} e^{2\alpha t_1}dt_1. \end{aligned}$$

Finally,

$$C_F(\tau) = \left(\frac{mk^2}{2\alpha}\right) e^{-\alpha\tau},$$

which is Eq. (4.2).

1. The joint density and the effective discount

We next solve the problem posed by Eqs. (4.10) and (4.11):

$$\frac{\partial \tilde{p}}{\partial t} = \left(\omega_1 - \alpha\omega_2 - i\frac{k^2}{2}\omega_2^2\right)\frac{\partial \tilde{p}}{\partial \omega_2} - i\alpha m\omega_2\tilde{p}, \tag{B5}$$

$$\tilde{p}(\omega_1,\omega_2,0|r_0) = e^{-i\omega_2 r_0}. \tag{B6}$$

Equation (B5) is a first-order partial differential equation which we solve by the method of characteristics [31]. The characteristic of Eq. (B5) is a function $h(\omega_2)$ such that its derivative is

$$\frac{1}{h'(\omega_2)} = \omega_1 - \alpha\omega_2 - i\frac{k^2}{2}\omega_2^2.$$

Hence,

$$h(\omega_2) = \frac{2}{\lambda(\omega_1)}\operatorname{arctanh}\left[\frac{\alpha + ik^2\omega_2}{\lambda(\omega_1)}\right], \tag{B7}$$

where

$$\lambda(\omega_1) = \sqrt{\alpha^2 + 2ik^2\omega_1}. \tag{B8}$$

Knowing the characteristic of Eq. (B5), the next step toward solving it consists of the change of variables:

$$\omega_2' = \omega_2, \quad t' = t + h(\omega_2), \tag{B9}$$

so

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t'}, \quad \frac{\partial}{\partial \omega_2} = \frac{\partial}{\partial \omega_2'} + h'(\omega_2)\frac{\partial}{\partial t'}.$$

In these variables Eq. (B5) reduces to (recall that $\omega_2' = \omega_2$)

$$\frac{1}{\tilde{p}}\frac{\partial \tilde{p}}{\partial \omega_2} = i\alpha m\omega_2 h'(\omega_2),$$

which can be integrated at once, yielding

$$\tilde{p}(\omega_1,\omega_2,t') = \phi(t')\exp\left[i\alpha m\int \omega_2 h'(\omega_2)d\omega_2\right],$$

where $\phi(t')$ is an arbitrary function to be determined by the initial condition (B6). Making use of the explicit expression of $h'(\omega_2)$ to perform the integration in the exponential and undoing the change of variables (B9), we write the general solution of Eq. (B5) as

$$\tilde{p}(\omega_1,\omega_2,t) = \frac{\phi[t + h(\omega_2)]e^{-\alpha\theta h(\omega_2)/2}}{\left(i\omega_1 - i\alpha\omega_2 + k^2\omega_2^2/2\right)^{\theta/2}}, \tag{B10}$$

where

$$\theta = \frac{2\alpha m}{k^2}. \tag{B11}$$

The initial condition (B6) implies that

$$e^{-i\omega_2 r_0} = \frac{\phi[h(\omega_2)]e^{-\alpha\theta h(\omega_2)/2}}{\left(i\omega_1 - i\alpha\omega_2 + k^2\omega_2^2/2\right)^{\theta/2}},$$

so

$$\phi[h(\omega_2)] = \left(i\omega_1 - i\alpha\omega_2 + \frac{k^2}{2}\omega_2^2\right)^{\theta/2} \times \exp\left[-i\omega_2 r_0 + \frac{\alpha}{2}\theta h(\omega_2)\right].$$

In this expression ϕ is written as a function of h and ω_2 (and, of course, t and ω_1) but h is related to ω_2 through Eq. (B7). Therefore, by inverting Eq. (B7) to write ω_2 in terms of h:

$$\omega_2 = i\frac{\alpha}{k^2} - i\frac{\lambda(\omega_1)}{k^2}\tanh[\lambda(\omega_1)h(\omega_2)/2],$$

we have

$$\phi(\xi) = \left\{i\omega_1 + \frac{\alpha}{k^2}\left[\alpha - \lambda(\omega_1)\tanh\frac{\lambda(\omega_1)\xi}{2}\right] - \frac{1}{2k^2}\left[\alpha - \lambda(\omega_1)\tanh\frac{\lambda(\omega_1)\xi}{2}\right]^2\right\}^{\theta/2} \times \exp\left\{\frac{1}{k^2}\left[\alpha - \lambda(\omega_1)\tanh\frac{\lambda(\omega_1)\xi}{2}\right]r_0 + \frac{\alpha}{2}\theta\xi\right\}.$$

Plugging this expression into the general solution (B10) we finally obtain the joint characteristic function

$$\tilde{p}(\omega_1,\omega_2,t|r_0) = \left[\frac{i\omega_1 + (\alpha/k^2)\varphi(t) - (1/2k^2)\varphi^2(t)}{i\omega_1 - i\alpha\omega_2 + k^2\omega_2^2/2}\right]^{\theta/2} \times \exp\left[\frac{1}{k^2}\varphi(t)r_0 + \frac{\theta}{2}\alpha t\right], \tag{B12}$$

where

$$\varphi(t) \equiv \alpha - \lambda(\omega_1)\tanh\left\{\tfrac{1}{2}\lambda(\omega_1)\left[t + h(\omega_2)\right]\right\}. \tag{B13}$$

We know discount is obtained from Eq. (B12) after setting $\omega_1 = -i$ and $\omega_2 = 0$. But [cf. Eqs. (B7)–(B8)]

$$\lambda(\omega_1 = -i) = \sqrt{\alpha^2 + 2k^2} \equiv \lambda \tag{B14}$$

and

$$h(0) = \frac{2}{\lambda}\operatorname{arctanh}(\alpha/\lambda). \tag{B15}$$

Hence, setting $\omega_1 = -i$ and $\omega_2 = 0$ in Eq. (B13), substituting for Eq. (B15) and taking into account Eq. (B14), we get

$$\varphi(t) = -\frac{2k^2\tanh(\lambda t/2)}{\lambda + \alpha\tanh(\lambda t/2)}. \tag{B16}$$

Discount will then be given by

$$D(t) = [1 + (\alpha/k^2)\varphi(t) - (1/2k^2)\varphi^2(t)]^{\theta/2} \times \exp\left[\frac{1}{k^2}\varphi(t)r_0 + \frac{\theta}{2}\alpha t\right].$$

Substituting for Eq. (B16) and some simple algebra finally yield

$$D(t) = \left[\frac{\lambda e^{\alpha t/2}}{\lambda\cosh(\lambda t/2) + \alpha\sinh(\lambda t/2)}\right]^{\theta} \times \exp\left[-\frac{2\sinh(\lambda t/2)}{\lambda\cosh(\lambda t/2) + \alpha\sinh(\lambda t/2)}\right], \tag{B17}$$

which is equivalent to the expression, Eq. (4.12), of the discount function given in the main text.

APPENDIX C: THE LOG-NORMAL PROCESS

The log-normal discount model is defined by the following pair of stochastic differential equations:

$$dx(t) = r(t)dt,$$
$$\frac{dr(t)}{r(t)} = \alpha dt + k dW(t).$$

The FPE for the joint PDF of the bidimensional process $(x(t),r(t))$ is [cf. Eq. (2.12)]

$$\frac{\partial p}{\partial t} = -r\frac{\partial p}{\partial x} - \alpha\frac{\partial}{\partial r}(rp) + \frac{1}{2}k^2\frac{\partial^2}{\partial r^2}(r^2 p), \tag{C1}$$

with the initial conditions given in Eq. (2.13). Let us denote by $\tilde{p}(\omega_1,\omega_2,t|r_0)$ the joint characteristic function (2.14). The joint Fourier transform of Eq. (C1) yields the following partial differential equation for $\tilde{p}$:

$$\frac{\partial \tilde{p}}{\partial t} = (\omega_1 + \alpha\omega_2)\frac{\partial \tilde{p}}{\partial \omega_2} + \frac{1}{2}k^2\omega_2^2\frac{\partial^2 \tilde{p}}{\partial \omega_2^2}, \tag{C2}$$

with initial condition

$$\tilde{p}(\omega_1,\omega_2,0|r_0) = e^{-i\omega_2 r_0}. \tag{C3}$$

We take, in addition to the Fourier transform with respect to x and r, the Laplace transform with respect to time:

$$\hat{q}(\omega_1,\omega_2,\sigma|r_0) = \int_0^\infty e^{-\sigma t}\tilde{p}(\omega_1,\omega_2,t|r_0)dt, \tag{C4}$$

and the initial-value problem (C2) and (C3) turns into the following ordinary differential equation:

$$\frac{1}{2}k^2\omega_2^2\frac{d^2\hat{q}}{d\omega_2^2} + (\omega_1 + \alpha\omega_2)\frac{d\hat{q}}{d\omega_2} - \sigma\hat{q} = -e^{-i\omega_2 r_0}. \tag{C5}$$

There are boundary conditions implicitly attached to this equation. In effect, let us note that $\hat{q}(\omega_1,\omega_2 = 0,\sigma|r_0) = \hat{q}(\omega_1,\sigma|r_0)$ is the marginal characteristic function of the process $x(t)$, since this distribution exists and is finite, we must have

$$\lim_{\omega_2\to 0}\hat{q}(\omega_1,\omega_2,\sigma|r_0) = \text{ finite}. \tag{C6}$$

On the other hand, for the inverse Fourier transform of $\hat{q}(\omega_1,\omega_2,\sigma|r_0)$ with respect to ω_2 to exist it is necessary that

$$\lim_{\omega_2\to\pm\infty}\hat{q}(\omega_1,\omega_2,\sigma|r_0) = 0. \tag{C7}$$

The change of the independent variable

$$\xi = \frac{2\omega_1}{k^2\omega_2}$$

turns Eq. (C5) into

$$\xi^2\frac{d^2\hat{q}}{d\xi^2} + \xi(\gamma - \xi)\frac{d\hat{q}}{d\xi} - \frac{2\sigma}{k^2}\hat{q} = -\frac{2}{k^2}e^{-2i\omega_1 r_0/k^2\xi}, \tag{C8}$$

where

$$\gamma = 2(1 - \alpha/k^2). \tag{C9}$$

Boundary conditions (C6) and (C7) now read

$$\lim_{\xi\to\infty}\hat{q}(\omega_1,\xi,\sigma|r_0) = \text{finite} \tag{C10}$$

and

$$\lim_{\xi\to 0}\hat{q}(\omega_1,\xi,\sigma|r_0) = 0. \tag{C11}$$

We define a new unknown function $\hat{\phi}(\omega_1,\xi,\sigma|r_0)$ related to $\hat{q}$ by

$$\hat{\phi} = \xi^{-\beta}\hat{q}, \tag{C12}$$

where β is an arbitrary exponent to be conveniently determined. Equation (C8) now reads

$$\xi^{\beta+2}\frac{d^2\hat{\phi}}{d\xi^2} + \xi^{\beta+1}(2\beta + \gamma - \xi)\frac{d\hat{\phi}}{d\xi} + \left\{\left[\beta(\beta-1) + \gamma\beta - \frac{2\sigma}{k^2}\right]\xi^\beta - \beta\xi^{\beta+1}\right\}\hat{\phi} = -\frac{2}{k^2}e^{-2i\omega_1 r_0/k^2\xi}. \tag{C13}$$

We choose the undetermined exponent $\beta = \beta(\sigma)$ such that

$$\beta(\beta-1) + \gamma\beta - \frac{2\sigma}{k^2} = 0,$$

that is [38],

$$\beta = \tfrac{1}{2}[1 - \gamma + \sqrt{(1-\gamma)^2 + 8\sigma/k^2}]. \tag{C14}$$

Then Eq. (C13) transforms into the following inhomogeneous Kummer equation:

$$\xi\frac{d^2\hat{\phi}}{d\xi^2} + (2\beta + \gamma - \xi)\frac{d\hat{\phi}}{d\xi} - \beta\hat{\phi} = -\frac{2/k^2}{\xi^{1+\beta}}e^{-2i\omega_1 r_0/k^2\xi}. \tag{C15}$$

The boundary conditions that accompany this equation are

$$\lim_{\xi\to\infty}[\xi^\beta\hat{\phi}(\omega_1,\xi,\sigma|r_0)] = \text{finite} \tag{C16}$$

and

$$\lim_{\xi\to 0}[\xi^\beta\hat{\phi}(\omega_1,\xi,\sigma|r_0)] = 0. \tag{C17}$$

Two independent solutions of the homogeneous Kummer equation, corresponding to the inhomogeneous Eq. (C15), are the confluent hypergeometric functions $F(\beta,2\beta+\gamma,\xi)$ and $U(\beta,2\beta+\gamma,\xi)$ of the first and second kinds, respectively [32]. These functions allow us to solve the inhomogeneous equation by the lengthy but otherwise standard method of variation of parameters. The solution obeying the boundary conditions (C16) and (C17) and written in the original variables $\hat{q}$ and ω_2 finally reads

$$\hat{q}(\omega_1,\omega_2,s|r_0) = \frac{2\Gamma(\beta)}{k^2\Gamma(2\beta+\gamma)}\left(\frac{2\omega_1}{k^2\omega_2}\right)^\beta\left[U\left(\beta,2\beta+\gamma,\frac{2\omega_1}{k^2\omega_2}\right)\right.$$

$$\times \int_0^{\frac{2\omega_1}{k^2\omega_2}} y^{\beta+\gamma-2} e^{-y-2i\omega_1 r_0/k^2 y} F(\beta,2\beta+\gamma,y)dy + F\left(\beta,2\beta+\gamma,\frac{2\omega_1}{k^2\omega_2}\right) \times \int_{\frac{2\omega_1}{k^2\omega_2}}^{\infty} y^{\beta+\gamma-2} e^{-y-2i\omega_1 r_0/k^2 y} U(\beta,2\beta+\gamma,y)dy\Bigg]. \tag{C18}$$

We know from Sec. II that the discount function $D(t)$ is obtained by setting $\omega_1 = -i$ and $\omega_2 = 0$ in the characteristic function $\tilde{p}(\omega_1,\omega_2,t|r_0)$ [cf. Eq. (2.15)]. For the log-normal model we know the Laplace transform of the characteristic function, $\hat{q}(\omega_1,\omega_2,\sigma|r_0)$, given in Eq. (C18) and whose analytical inversion yielding $\tilde{p}$ seems to be beyond reach. We, therefore, alternatively obtain the Laplace transform of the discount function,

$$\hat{D}(\sigma) = \int_0^{\infty} e^{-\sigma t} D(t)dt,$$

which, in terms of $\hat{q}$, is given by

$$\hat{D}(\sigma) = \hat{q}\big(\omega_1 = -i,\omega_2 = 0,\sigma|r_0\big).$$

Setting $\omega_1 = -i$ in Eq. (C18), taking the limit $\omega_2 \to 0$, and bearing in mind the following property of the Kummer function U [32]:

$$\lim_{z\to\infty}[z^a U(a,c,z)] = 1,$$

we finally get

$$\hat{D}(\sigma) = \frac{2\Gamma(\beta)}{k^2\Gamma(2\beta+\gamma)} \times \int_0^{\infty} \frac{e^{-(2r_0\zeta/k^2+1/\zeta)}}{\zeta^{\beta+\gamma}} F(\beta,2\beta+\gamma,\zeta^{-1})d\zeta. \tag{C19}$$

1. Asymptotic expressions

In order to find asymptotic expressions of $D(t)$ for large values of t we will use the so-called Tauberian theorems which under rather general conditions relate the long-time behavior of any function with the small σ behavior of its Laplace transform [33]. Looking at Eq. (C19) we notice that the σ dependence is through the quantity $\beta = \beta(\sigma)$ defined in Eq. (C14). We, therefore, assume σ small and expand $\beta(\sigma)$ up to first order:

$$\beta = \frac{1}{2}\left[1-\gamma+|1-\gamma|+\frac{4\sigma}{k^2|1-\gamma|}+O(\sigma^2)\right]. \tag{C20}$$

As this expansion shows, the approximate form of $\hat{D}(s)$ as $\sigma \to 0$—and, hence, the asymptotic form of $D(t)$ as $t \to \infty$—will depend on the range of values taken by the dimensionless parameter γ which, in turn, depends on the "signal-to-noise ratio" α/k^2 [cf. Eq. (C9)]. We single out two regions, $\gamma > 1$ and $\gamma < 1$, separated by the value $\gamma = 1$. In each case discounting exhibits a markedly distinct behavior as time progresses. Let us note that $\gamma > 1$ corresponds to a signal weaker than noise ($\alpha < k^2/2$) and $\gamma < 1$ to a signal stronger than noise ($\alpha > k^2/2$), while the limit case $\gamma = 1$ represents a sort of equilibrium situation in which signal equals noise ($\alpha = k^2/2$).

(i) When $\gamma > 1$ ($\alpha < k^2/2$) we have $|1-\gamma| = \gamma - 1$ and

$$\beta = \frac{2\sigma}{k^2(\gamma-1)} + O(\sigma^2). \tag{C21}$$

From the definition of the Kummer function [32]

$$F(a,c,z) = 1 + \frac{a}{c}\frac{z}{1!} + \frac{a(a-1)}{c(c-1)}\frac{z^2}{2!} + \cdots, \tag{C22}$$

and taking into account Eq. (C21), we get

$$F(\beta,2\beta+\gamma,\xi^{-1}) = 1 + O(\sigma)$$

and

$$\xi^{-\gamma-\beta} = \xi^{-\gamma}[1+O(\sigma)].$$

Since $\Gamma(z) = 1/z + O(1)$ and $\Gamma(z+a) = \Gamma(a) + O(z)$ as $z \to 0$ [32] we write [cf. Eq. (C21)]

$$\frac{\Gamma(\beta)}{\Gamma(2\beta+\gamma)} = \frac{1}{\beta\Gamma(\gamma)}[1+O(\beta)] = \frac{k^2(\gamma-1)}{2\sigma\Gamma(\gamma)}[1+O(\sigma)].$$

Collecting results into Eq. (C19) yields

$$\hat{D}(\sigma) = K_1(r_0)\frac{1}{\sigma} + O(1),$$

where

$$K_1(r_0) = \frac{\gamma-1}{\Gamma(\gamma)}\int_0^{\infty} \zeta^{-\gamma} e^{-(2r_0\zeta/k^2+1/\zeta)}d\zeta. \tag{C23}$$

Finally, recalling the standard property of the Laplace transform

$$\lim_{t\to\infty} D(t) = \lim_{\sigma\to 0}[\sigma\hat{D}(\sigma)],$$

we conclude that when the signal is weaker than noise, discount saturates towards a constant value

$$\lim_{t\to\infty} D(t) = K_1(r_0) = \text{ const.} \tag{C24}$$

(ii) When $\gamma < 1$ ($\alpha > k^2/2$) we have $|1-\gamma| = 1-\gamma$ and expansion (C20) now reads

$$\beta = 1-\gamma+\frac{2\sigma}{k^2(1-\gamma)} + O(\sigma^2). \tag{C25}$$

Hence $\gamma+\beta = 1+O(\sigma)$, whence

$$\zeta^{\gamma+\beta} = \zeta[1+O(\sigma)]. \tag{C26}$$

Also, from Eq. (C22) we have

$$F(\beta,2\beta+\gamma,\zeta^{-1}) = F(1-\gamma,2-\gamma,\zeta^{-1}) + O(\sigma). \tag{C27}$$

On the other hand, the Taylor expansion of $\Gamma(\beta)$ around $\sigma = 0$ yields

$$\Gamma(\beta) = \Gamma(1-\gamma)\left[1 + \frac{2\psi(1-\gamma)}{k^2(1-\gamma)}\sigma + O(\sigma^2)\right],$$

where

$$\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)} = \frac{d}{dz}\ln\Gamma(z)$$

is the digamma function. Analogously,

$$\Gamma(2\beta+\gamma) = \Gamma(2-\gamma)\left[1 + \frac{4\psi(2-\gamma)}{k^2(1-\gamma)}\sigma + O(\sigma^2)\right].$$

Therefore,

$$\begin{aligned}
&\frac{\Gamma(\beta)}{\Gamma(2\beta+\gamma)} \\
&= \frac{\Gamma(1-\gamma)}{\Gamma(2-\gamma)}\,\frac{1 + \frac{2\psi(1-\gamma)}{k^2(1-\gamma)}\sigma + O(\sigma^2)}{1 + \frac{4\psi(2-\gamma)}{k^2(1-\gamma)}\sigma + O(\sigma^2)} \\
&= \frac{1/(1-\gamma)}{\left[1 + \frac{4\psi(2-\gamma)}{k^2(1-\gamma)}\sigma + O(\sigma^2)\right]\left[1 - \frac{2\psi(1-\gamma)}{k^2(1-\gamma)}\sigma + O(\sigma^2)\right]} \\
&= \frac{1/(1-\gamma)}{1 + \frac{2}{k^2(1-\gamma)}[2\psi(2-\gamma) - \psi(1-\gamma)]\sigma + O(\sigma^2)} \\
&= \frac{1}{1-\gamma+(2\delta/k^2)\sigma + O(\sigma^2)},
\end{aligned}$$

where

$$\delta = 2\psi(2-\gamma) - \psi(1-\gamma). \tag{C28}$$

We thus write

$$\frac{\Gamma(\beta)}{\Gamma(2\beta+\gamma)} = \frac{k^2/2\delta}{\varrho + \sigma + O(\sigma^2)}, \tag{C29}$$

where

$$\varrho = \frac{k^2(1-\gamma)}{2\delta}, \tag{C30}$$

which, after substituting for Eq. (C9), yields

$$\varrho = \frac{1}{\delta}\left(\alpha - \frac{k^2}{2}\right). \tag{C31}$$

Collecting (C26)–(C29) into Eq. (C19) we get

$$\hat{D}(\sigma) = \frac{K_2(r_0)}{\varrho + \sigma}[1 + O(\sigma)], \tag{C32}$$

where

$$K_2(r_0) = \frac{1}{\delta}\int_0^\infty \zeta^{-1} e^{-(2r_0\zeta/k^2 + 1/\zeta)} F(1-\gamma, 2-\gamma, \zeta^{-1})\, d\zeta. \tag{C33}$$

Tauberian theorems [33] allow us to obtain the asymptotic long-time behavior of $D(t)$ by means of the Laplace inversion of the approximate expression (C32). This yields the exponential decay:

$$D(t) \sim K_2(r_0) e^{-\varrho t} \tag{C34}$$

($t \to \infty$).

Let us finally prove that $\varrho > 0$. To this end, looking at Eq. (C30) and bearing in mind that $\gamma < 1$, we should prove that $\delta > 0$. In effect, using the property [32]

$$\psi(1+x) = \psi(x) + \frac{1}{x}, \tag{C35}$$

we see that $\psi(1+x) > \psi(x)$ if $x > 0$, which implies $\psi(2-\gamma) > \psi(1-\gamma)$ as long as $\gamma < 1$. From Eq. (C28) it immediately follows that $\delta > 0$ and hence $\varrho > 0$. Moreover, from the definition of γ given in Eq. (C9) and the property (C35) we write the following more convenient form of the dimensionless parameter δ:

$$\delta = \psi(2\alpha/k^2) + \frac{1}{2\alpha/k^2 - 1}, \tag{C36}$$

and we recall that $\gamma < 1$ implies that $2\alpha/k^2 > 1$.

(iii) When $\gamma = 1$ ($\alpha = k^2/2$) we need not expand $\beta(\sigma)$ in powers of σ, as we did in Eq. (C20), because form the definition (C14) we get the exact expression

$$\beta = \frac{\sqrt{2\sigma}}{k},$$

allowing us to write

$$\begin{aligned}
\frac{\Gamma(\beta)}{\Gamma(2\beta+\gamma)} &= \frac{k}{\sqrt{2\sigma}}[1 + O(\sqrt{\sigma})], \\
\zeta^{-\gamma-\beta} &= \zeta^{-1}[1 + O(\sqrt{\sigma})],
\end{aligned}$$

and

$$\begin{aligned}
F(\beta, 2\beta+\gamma, \zeta^{-1}) &= F(0,1,\zeta^{-1}) + O(\sqrt{\sigma}) \\
&= 1 + O(\sqrt{\sigma}).
\end{aligned}$$

Collecting results we have

$$\hat{D}(\sigma) = \frac{\sqrt{2}K_3(r_0)}{k\sqrt{\sigma}}[1 + O(\sqrt{\sigma})],$$

where

$$K_3(r_0) = \int_0^\infty \zeta^{-1} e^{-(2r_0\zeta/k^2 + 1/\zeta)} d\zeta.$$

Tauberian theorems tell us that the long-time behavior of $D(t)$ will be given by the Laplace inversion of $\hat{D}(\sigma) \sim 1/\sqrt{\sigma}$ ($\sigma \to 0$). That is,

$$D(t) \sim \left(\frac{2}{\pi}\right)^{1/2} \frac{K_3(r_0)}{k\sqrt{t}} \tag{C37}$$

($t \to \infty$), which is the hyperbolic discount previously obtained by two of us [19,20].

[1] P. Samuelson, A note on measurement of utility, Rev. Econ. Stud. **4**, 155 (1937).

[2] P. Dasgupta, *Human Well-Being and the Natural Environment* (Oxford University Press, Oxford, 2004).

[3] K. J. Arrow *et al.*, Determining benefits and costs for future generations, Science **341**, 349 (2013).

[4] N. Stern, *The Economics of Climate Change: The Stern Review* (Cambridge University Press, Cambridge, UK 2006).

[5] W. D. Nordhaus, The Stern review on the economics of climate change, J. Econ. Lit. **45**, 686 (2007).

[6] W. D. Nordhaus, Critical assumptions in the Stern review on climate change, Science **317**, 201 (2007).

[7] N. Stern, Ethics, equity and the economics of climate change. Paper 1: Science and philosophy, Econ. Philos. **30**, 397 (2014).

[8] N. Stern, Ethics, equity and the economics of climate change. Paper 2: Economics and politics, Econ. Philos. **30**, 445 (2014).

[9] G. M. Heal and A. Millner, Agreeing to disagree on climate policy, Proc. Natl. Acad. Sci. USA **111**, 3695 (2014).

[10] F. P. Ramsey, A. mathematical theory of saving, Econ. J. **38**, 543 (1928).

[11] O. Vasicek, An equilibrium characterization of the terms structure, J. Fin. Econ. **5**, 177 (1977).

[12] J. C. Cox, J. E. Ingersoll, and S. A. Ross, A. theory of the term structure of interest rates, Econometrica **53**, 385 (1985).

[13] L. U. Dothan, On the term structure of interest rates, J. Finan. Econ. **6**, 59 (1978).

[14] D. Brigo and F. Mercurio, *Interest Rate Models—Theory and Practice* (Springer-Verlag, Berlin, 2006).

[15] L. B. G. Andersen and V. V. Piterberg, *Interest Rate Modeling*, Vols. I—III (Atlantic Financial Press, Boston, MA, 2010).

[16] D. Duffie, Credit risk modeling with affine processes, J. Bank, Finan, **29**, 2751 (2005).

[17] R. Newell and N. Pizer, Discounting the distant future: How much do uncertain rates increase valuations?, J. Environ. Econ. Manag. **46**, 52 (2003).

[18] J. D. Farmer, J. Geanakoplos, J. Masoliver, M. Montero, and J. Perelló, Discounting the distant future, Cowles Foundation Discussion Paper No. 1951, 2014, Available at SSRN: http://ssrn.com/abstract=2465953.

[19] J. D. Farmer and J. Geanakoplos, Hyperbolic discounting is rational: Valuing the far future with uncertain discount rates, Cowles Foundation Discussion Paper No. 1719, 2009, Available at SSRN: http://ssrn.com/abstract=1448811.

[20] J. Geanakoplos, W. Sudderth, and O. Zeitouini, Asymptotic behavior of stochastic discount rate, Ind. J. Stat. **76 A**, 150 (2014).

[21] G. E. Uhlenbeck and L. S. Ornstein, On the theory of Brownian motion, Phys. Rev. **36**, 823 (1930).

[22] W. Feller, Two singular diffusion processes, Ann. Math. **54**, 173 (1951).

[23] M. F. M. Osborne, Brownian motion in the stock market, Oper. Res. **7**, 145 (1959), reprinted in P. H. Cootner (ed.) *The Random Character of Stock Market Prices* (M.I.T. Press, Cambridge, MA, 1964).

[24] J. C. Cox, J. E. Ingersoll, and S. A. Ross, A re-examination of traditional hypothesis about the term structure of interest rates, J. Finan. **36**, 769 (1981).

[25] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1985).

[26] M. Kac, On the distribution of certain Wiener functionals, Trans. Am. Math Soc. **65**, 1 (1949).

[27] B. Stehlíková and D. Ševčovič, On the non-existence of a one-factor interest rate model for volatility averaged generalized Fong-Vasicek term structures, *Proceedings of the Czech-Japanese Seminar in Applied Mathematics*, edited by M. Beneŝ, M. Kimura, T. Nakaki (Takachiho/University of Miyazaki, Japan, 2008), p. 40, Available at SSRN: http://ssrn.com/abstract=1295557.

[28] Y. K. Kwok, *Mathematical Models of Financial Derivatives*, 2nd ed. (Springer-Verlag, Berlin, 2008).

[29] B. E. Baaquie, *Quantum Finance* (Cambridge University Press, Cambridge, UK, 2004).

[30] J. Masoliver and J. Perelló, First-passage and escape problems in the Feller process, Phys. Rev. E **86**, 041116 (2012).

[31] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2 (J. Wiley-VCH, New York, 1989).

[32] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, Berlin, 1966).

[33] R. A. Handelsman and J. S. Lew, Asymptotic expansion of Laplace convolutions for large argument and tail densities for certain sums of random variables, SIAM J. Math. Anal. **5**, 425 (1974).

[34] See, for instance, Refs. [14–16] and references therein.

[35] M. L. Weitzman, Why the far-distant future should be discounted at its lowest possible rate?, J. Environ. Econ. Manag. **36**, 201 (1998).

[36] C. Gollier, P. Koundouri, and T. Pantelidis, Declining discount rates: economic justifications and implications for the long-run policy, Econ. Pol. **23**, 757 (2008).

[37] S. Giglio, M. Maggiori, and J. Stroebel, Very long-run discount rates, Quarterly Journal of Economics **130**, 1 (2015).

[38] Out of the two possible solutions for β, we discard the one with a minus sign in front of the square root because the corresponding $\hat{\phi}$ does not satisfy boundary conditions.