DISCOUNTING THE DISTANT FUTURE

By

J. Doyne Farmer, John Geanakoplos, Jaume Masoliver,
Miquel Montero, and Josep Perelló

July 2014

COWLES FOUNDATION DISCUSSION PAPER NO. 1951

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Discounting the Distant Future

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

[1]Mathematical Institute and Institute for New Economic Thinking, Oxford Martin School, University of Oxford.
[2]Department of Economics, Yale University, New Haven CT.
[3]Santa Fe Institute, Santa Fe, New Mexico.
[4]Departament de Física Fonamental, Universitat de Barcelona.

May 27, 2014

Abstract

If the historical average annual real interest rate is $m > 0$, and if the world is stationary, should consumption in the distant future be discounted at the rate of m per year? Suppose the annual real interest rate $r(t)$ reverts to m according to the Ornstein Uhlenbeck (OU) continuous time process $dr(t) = \alpha\big[m - r(t)\big]dt + kdw(t)$, where w is a standard Wiener process. Then we prove that the long run rate of interest is $r_\infty = m - k^2/2\alpha^2$. This confirms the Weitzman-Gollier principle that the volatility and the persistence of interest rates lower long run discounting. We fit the OU model to historical data across 14 countries covering 87 to 318 years and estimate the average short rate m and the long run rate r_∞ for each country. The data corroborate that, when doing cost benefit analysis, the long run rate of discount should be taken to be substantially less than the average short run rate observed over a very long history.

JEL classification: C1, G12, Q5.

Key words: Discounting, environment, interest rates, inflation, Ornstein-Uhlenbeck process.

I. Introduction

For environmental problems such as global warming, future costs must be balanced against present costs. This is traditionally done by using an exponential discount function with a constant discount rate, usually taken to be equal to the historical mean real interest rate. For nine countries with stable rates, including the United States and England, we estimate that the historical mean real interest rate is nearly 3% per annum, which implies a negligible importance to events out several hundred years.

The choice of discount rate has generated a major controversy as to the urgency for immediate action. At very low discount rates it makes sense to expend resources today to stave off environmental disasters two centuries down the road. At high enough interest rates, costly action today for the same purpose would appear to be foolish. The choice of a discounting function has enormous consequences for long run environmental planning (Dasgupta, 2004). For example, in a highly influential report on climate change commissioned by the UK government, Stern (2006) 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 (2007b) argues for a discount rate of 4%, which implies a present value of 2%, and at other times has advocated rates as high as 6% (Nordhaus, 2007a), 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 (Arrow et al., 2013). Stern has been widely criticized for using such a low rate (Nordhaus, 2007b,a; Dasgupta, 2006; Mendelsohn, 2006; Weitzman, 2007; Nordhaus, 2008). This issue is likely to surface again with the upcoming Calderon report in July 2014.

The normative approach to choosing the discount rate attempts to derive the right discount from axiomatic principles of justice, or from utility theory and assumptions about growth. Economists present a variety of reasons for discounting, including impatience, economic growth, and declining marginal utility; these are embedded in the Ramsey formula, which forms the basis for one standard approach to discounting the distant future (Arrow et al., 2012).

A positive approach attempts to ascertain 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 (which can also sometimes be deduced from market prices) one can infer the market discount rate for real consumption. Unfortunately, for horizons much beyond 30 years, market data becomes thin.

Faced with this situation, the practical positive economist who wants to engage in environmental policy debates (and who therefore cannot say the future is unknowable) would naturally be tempted to say "over the last two hundred years real interest rates have averaged 3% per annum, so let us use 3% as the discount rate moving forward", or he might say "in Wall Street's forward looking models, the average real interest rate over the next 30 years is also about 3%, so let's use 3% as the discount moving forward". We argue to the contrary, that the practical economist should deduce from historical fluctuations of short real interest rates that it is appropriate to use a discount rate considerably below the average interest rates in his models. We do this by taking the two most basic stationary stochastic interest rate models used on Wall Street and proving that if they are used to price bonds of arbitrarily long maturity, the yield will converge to a number which is well below the average short term rate generated by the model through time. Our method of analysis is to take continuous time limits of these standard models, expressing the pricing formulas as partial differential equations, and then doing Fourier analysis.

The idea that fluctuations in economic performance cannot be ignored, because volatile real rates that are persistent can lead to long run real rates that are lower than the mean, has emerged in a growing body of work. The pioneering papers of Weitzman (1998) and Gollier et al. (2008) are based on the mathematical analysis of an extreme stylized case in which future annual real rates are unknown today, but starting tomorrow at $t = 1$ will be fixed forever (i.e. be completely persistent) at one of a finite number of values. They argued that in such a situation the appropriate long-run interest rate at $t = 0$ is the lowest rate there could be at $t = 1$. Of course their environment is far from stationary, depending as it does on a once and for all permanent change, after which the world does not resemble its history. Dybvig, Ingersoll, and Ross (1996) had previously proved with great generality (in particular, with or without persistence), that as long as there is no arbitrage, the long run rate of interest at any date t, like $t = 0$, must be less or equal than the long run rate of interest at any future time like $t = 1$. Thus any increase in uncertainty about the future long run rate of interest implies a lower long run rate of interest today. The Weitzman and Gollier model is just a special case. It is therefore not clear from the Weitzman and Gollier example that persistence (rather than just volatility) really is a central factor in lowering the long term interest rate. Nonetheless, we shall show that in our stationary models, volatility and persistence are indeed the drivers of lower long term interest rates, confirming the Weitzman-Gollier intuition.

Litterman, Scheinkman, and Weiss (1991), Newell and Pizer (2003), and Groom et al. (2007) simulated more realistic stochastic interest rate processes than Weitzman and Gollier, out to horizons of a few hundred years, leaving aside the asymptotic (infinite horizon) behavior of real rates. They found indeed that as the horizon gets longer, the long run rate of interest tends to get lower. Groom et al (2007) noted that the drop in rates does seem to depend on volatility and persistence, depending also on the particular model and parameters.

Farmer and Geanakoplos (2009) used the reflection principle to prove that when the interest rate r(t) follows a geometric random walk (as in Litterman et al. 1991 and Newell and Pizer 2003 simulations), the price $D(t)$ at time 0 for one unit of consumption at time t is approximately $K/\sqrt{t}$ for all large t, where K is a constant. They called this hyperbolic discounting because the discount factor $D(t)$ is hyperbolic. The hyperbolic $D(t)$ is eventually substantially greater than any exponentially decaying function, showing that there is no positive long run rate of interest in the geometric random walk model. The long run rate of interest is 0, but that does not convey as precise information as saying $D(t)$ is approximately $K/\sqrt{t}$ for all large t. Since the sum of all these $D(t)$ is infinite, such $D(t)$ assign infinite value to any permanent positive flow of consumption: the infinite future is infinitely valuable.

By taking the continuous time limit of the geometric random walk model, and then examining the appropriate partial differential equation, we reproduce and extend the hyperbolic result of Farmer and Geanakoplos (2009). In that model real interest rates diverge from their starting point with higher and higher probability, and never go negative. In the geometric random walk, the long run rate of interest is below all possible short run interest rates. There is evidence, however, that real rates display mean reversion (see Freeman et al.

(2013)), and there have been many epochs in the United States with negative real rates of interest (including the period in which this paper was written).

We are therefore led (by tractability as well) to consider another one of the standard models of finance, in which the real interest rate follows the Ornstein-Uhlenbeck (OU) continuous time process

$$dr(t) = \alpha\big[m - r(t)\big]dt + kdw(t),$$

where w is a Wiener process. In this model real interest rates can go arbitrarily low (even arbitrarily negative) but they are always pulled back to $m > 0$. Is there a well defined long run rate of interest? If so, do the volatility and persistence of interest rates $r(t)$ affect the long rate? Can the economist infer what this long rate is from data about the stationary distribution of $r(t)$, including its mean m and its standard deviation?

We are able to give precise answers to these questions. We show in Section III that the joint probability density of the rate $r(t)$ and an auxiliary quantity $x(t)$ (defined through $r(t)$ as $dx(t)/dt = r(t)$) obeys a Fokker-Planck equation. By taking the Fourier transform of this density we are able to derive an exact formula for $D(t)$ and to prove that the long run rate of interest exists and is

$$r_\infty = m - k^2/2\alpha^2.$$

This is positive when k and $1/\alpha$ are both small. Real rates can become unboundedly negative; for any level there is a positive probability that at some point $r(t)$ becomes smaller. Yet the long run rate can remain positive. Most importantly, the long run rate of interest is below the average short run rate of interest.

The parameter k denotes the local volatility of the interest rate process, and the parameter α designates how fast $r(t)$ reverts to m; its reciprocal $1/\alpha$ indicates how long deviations from m persist. Thus our formula confirms the Weitzman-Gollier principle that the volatility k and the persistence $1/\alpha$ of interest rates lower long run discounting.

More surprisingly, it shows that knowledge of the stationary distribution of the interest rates $r(t)$ can be completely misleading about the long run rate of interest r_∞. The stationary distribution of $r(t)$ is well known to be normal, with mean m and variance $k^2/2\alpha$. By choosing k and α appropriately, we can make the stationary distribution of $r(t)$ as tight as we like around m, while at the same time making the long run rate r_∞ as far below m as we like. This shows the folly of discounting the far future on historical measures of average annual interest rates, even if there have been few large deviations from the average. To the best of our knowledge, we are the first to derive the long run rate of interest for this model.[1]

We try to get a sense of how practically important our results are by fitting our Ornstein Uhlenbeck model to historical data across 14 countries, and then computing the difference between m and r_∞ in the fitted models. In Section IV we estimate the parameters m, a, and k separately for each of 14 countries from data covering between 87 to 318 years, and then

[1] The closest other work of which we are aware is that of Davidson, Song, and Tippett (2013), who examined a square root Ornstein Uhlenbeck model in which $dx(t) = -\alpha xt)dt + kdw(t)$ and $r(t) = x^2(t)$. In the square root Ornstein Uhlenbeck process interest rates can never go negative, and they are pulled down toward 0, which is an absorbing state. Davidson et al. (2013) show that despite the fact that interest rates tend to drift toward zero, the expected short interest rate is positive and greater than the long run rate. They also solve for the long rate by studying a partial differential equation using the Feynman-Kac functional, which is quite different from our approach. As we will demonstrate from our empirical work, the failure to allow negative interest rates does not conform with historical data.

we plug these parameters into our formula for r_∞. We find that for five of the countries, r_∞ is actually negative. The nine countries with positive r_∞ include the US, England, and Argentina, but surprisingly this does not include Germany, Italy or Japan, due to epochs of runaway inflation. For the nine countries with positive r_∞ we find a modest but significant drop on average of 25% from m to r_∞. In the United States, for example, $m = 2.6\%$ and r_∞ is 2.1%. In Argentina, $m = 2.4\%$ and r_∞ is 1.1%. Spain is an extrene example with a positive short term rate of $m = 5.7\%$ but a negative long run rate $r_\infty = -6.4\%$. The data confirm, to the extent that they are reliable, that when doing cost benefit analysis, the long run rate of discount should be taken to be substantially less than the average short run rate observed over a very long history.

II. The discrete model

Consider first a discrete version of our model in which time and uncertainty are represented by a binary tree S, with a root 0 and such that every node $s \in S$ has exactly two immediate successors $s' = s_u$ and $s'' = s_d$. We denote the time $\tau(s)$ of any node s by the length of the path $(0, s]$ from the root 0 to s, not including 0. At each node we imagine a single consumption good. We assume that every node s is associated with a one period real interest rate $r(s)$, $-\infty < r(s) < \infty$.

A security is defined by the promise of the consumption good it makes in each node s. For us the most important securities are the riskless period t bonds that promise one unit of the consumption good in every state s with $\tau(s) = t$. We assume that every period t bond is traded at a price $D_s(t)$ at each node s. We assume that $D_s(t) = 0$ if $\tau(s) > t$, $D_s(t) = 1$ if $\tau(s) = t$, and that $D_s(t) = e^{-r(s)}$ if $\tau(s) = t - 1$.

We then define the discount factors as $D(t) \equiv D_0(t)$ and the key question is what can we say about $D(t)$ as $t \to \infty$? Sometimes it is convenient to express $D(t)$ in terms of the yield $y(t) = -(1/t) \ln D(t)$. The yield is the equivalent rate of interest that would explain the price:

$$D(t) = e^{-ty(t)}.$$

If the $y(t)$ converge to some y then we say that y is the long run rate of interest.

Suppose the security prices are such that there is "no arbitrage", that is, suppose there is no way to trade securities at the given prices across nodes in the tree so as to make a profit at some node without losing any money at any other node. Then the fundamental theorem of finance tells us that there must be probabilities $\gamma_{s_u} > 0, \gamma_{s_d} > 0$ with $\gamma_{s_u} + \gamma_{s_d} = 1$, such that for all nodes s and all bonds $t > \tau(s)$

$$D_s(t) = e^{-r(s)}[\gamma_{s_u} D_{s_u}(t) + \gamma_{s_d} D_{s_d}(t)].$$

Iterating this formula backward, we see that

$$D(t) \equiv D_0(t) = \mathrm{E}_\gamma \left[\exp \left(- \sum_{t'=0}^{t-1} r(t') \right) \right], \tag{1}$$

where the expectation $\mathrm{E}_\gamma[\cdot]$ is an average with respect to the probabilities γ over all possible interest rate paths that begin at the root 0 and terminate at a node with time $t - 1$.

The assignment $r(s)$ for each node s and the transition probabilities γ determine the price of every security from the above formula. The prices $D(t)$ can be computed by brute force. But since the number of paths grows exponentially in t, brute force enables us to compute bond prices exactly for bounded t. This horizon can be pushed much further out by Monte Carlo sampling of the paths, but even that will have computational limits.

Rather than computing the $D(t)$ by simulation, we seek to determine $D(t)$ for arbitrarily large t by analytical methods. Of course this will only be possible for tractable specifications of r and γ. One must then inquire how realistic those models are. [2]

While working together on Wall Street, Litterman, Scheinkman and Weiss (1991) investigated the geometric random walk model where

$$r(0) = r_0, \quad r(s_u) = vr(s), \quad \text{and} \quad r(s_d) = r(s)/v$$

for all $s \in S$, and

$$\gamma_{s_u} = \gamma_{s_d} = 1/2.$$

They simulated the model out for several hundred periods, finding that the yield $y(t)$ first rose and then fell toward 0. Newall and Pfizer (2003) reached similar conclusions in their simulations of the same interest rate process.

By using the Reflection Principle of the symmetric random walk, Farmer and Geanakoplos (2009) were able to prove that for any $r_0 > 0$ and any $v > 1$, for large t,

$$D(t) \sim 1/\sqrt{t}.$$

This confirms that the yield $y(t)$ converges to zero, but gives more information. It gives a rational justification to hyperbolic discounting. Geanakoplos, Sudderth, and Zeitouni (2014) generalize the same result to models in which the interest rate tree can have more than two branches, provided that the geometric average of the one year rates remains constant at r_0.

Another way to specify the geometric random walk is to work with the auxiliary variable $\xi(s)$, where

$$\xi(0) = \xi_0, \qquad \xi(s_u) = \xi(s) + k \qquad \text{and} \qquad \xi(s_d) = \xi(s) - k$$

for all $s \in S$, and $\gamma_{s_u} = \gamma_{s_d} = 1/2$. Defining $r(s) = e^{\xi(s)}$ gives the geometric random walk, which is also called the log normal model. Letting $r(s) = \xi(s)$ we get the normal random walk. More generally, we can define the normal random walk with mean reversion by

$$r(s_u) = r(s) + \alpha[m - r(s)] + k \qquad \text{and} \qquad r(s_d) = r(s) + \alpha[m - r(s)] - k$$

for all $s \in S$, and $\gamma_{s_u} = \gamma_{s_d} = 1/2$. The log normal random walk and the normal random walk with mean reversion are the two most basic stochastic interest rate models in finance.

[2]It is worth observing that simple though the binomial model appears, it can be used as a building block for much more complicated and realistic interest rate processes. For example, suppose that we extend the state space S to $Z \times S$ but assume $r(z, s) = r(s)$, in order to allow for signals z about which path future interest rates will take. The transition probabilities γ would then specify the probabilities of every (z', s') conditional on each (z, s): the simple transition probability γ_{s_u} is modified by the presence of z to γ_{zs_u}. It may well be that z provides information about which path in S will be taken over the next few time steps, but no information in the long run about $r(s')$. In that case the long run discount rate r_∞ can be derived from the simpler underlying binomial model.

In the next section we analyze the continuous limit of these two interest rate processes. By using Fourier-transform methods we are able to deduce the asymptotic behavior of $D(t)$ for both models, and to derive exact expressions for D(t) for all t in the Ornstein-Uhlenbeck model.

III. The Long Run Rate of Interest

We now present more general and more powerful results that can be obtained via the use of continuous time models. To understand how discounting depends on the random process used to characterize interest rates, we have studied the Ornstein-Uhlenbeck (OU) process (Uhlenbeck and Ornstein, 1930) and the log normal process (Osborne, 1959), using both analytic and numerical methods. Both of these are widely used in the interest rate pricing literature (Jouini, Cvitanic, and Musiela, 2001). In the log normal model rates cannot take negative values while in the OU model $r(t)$ can be either positive or negative. Another difference is that in the OU model rates are mean reverting while in the log normal model they are not.

A. The Ornstein-Uhlenbeck model

We first focus our attention on the Ornstein-Uhlenbeck model, which allows negative interest rates. The model is the continuous limit of the ordinary random walk with mean reversion described above. It can be defined through the stochastic differential equation

$$dr(t) = -\alpha\big[r(t) - m\big]dt + k dw(t), \tag{2}$$

where $r(t)$ is the real interest rate and $w(t)$ is the Wiener process. The parameter m is a mean value to which the process reverts, k is the amplitude of fluctuations, and α is the strength of the reversion to the mean.

Letting $r_0 = r(0)$ be the initial return, the probability density function $p(r, t|r_0)$ [3] is a normal distribution, which in the large time limit has mean m and variance

$$\sigma^2 = k^2/2\alpha. \tag{3}$$

The OU process has a stationary normal distribution with mean and standard deviation (m, σ).

Using Fourier transform methods described below, we derive an exact solution for the discount function $D(t)$ of the time-dependent OU model. We prove that in the limit $t \to \infty$ the discount function decays exponentially, i.e.

$$D(t) \simeq e^{-r_\infty t}, \tag{4}$$

where

$$r_\infty = m - k^2/2\alpha^2. \tag{5}$$

[3]The probability distribution was first obtained by G. E. Uhlenbeck and L. S. Ornstein in 1930 (Uhlenbeck and Ornstein, 1930). In the Appendix A we present an alternative derivation of $p(r, t|r_0)$ within the context of the present work.

Thus the long-run interest rate r_∞ is always lower than the average interest rate m, by an amount that depends on the noise parameter k and the reversion parameter α. From equations (3) and (5) it is evident that for any given mean interest rate m, by varying k and α, the long-run discount rate r_∞ can take on any value less than m, including negative values, while at the same time the standard deviation σ can also be made to take on any arbitrary positive value. In particular, by choosing the appropriate (k, α), we can make r_∞ arbitrarily far below m and σ arbitrarily small. The probability that $r(t) < r_\infty$ can be arbitrarily small, even when $r_\infty \ll m$ (see Appendix A). Deducing (perhaps from a long historical data base) the correct parameters (m, σ) of the stationary distribution of short run interest rates does not determine r_∞ by itself; on the contrary, any $r_\infty < m$ is consistent with them. To infer r_∞ from the data one must also tease out the mean reversion parameter α. Holding the long run distribution (m, σ) constant, by raising the persistence parameter $1/\alpha$ it is possible to lower r_∞ to any desired level. On the other hand, we also see from Eq. (5) that the long-run interest rate may be negative. How is it possible for r_∞ 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. (6) below 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 discounting rate can be much lower than the mean, and indeed can correspond to low interest rates that are rarely observed. This dramatically illustrates the folly of assuming that the average one period real interest rate is the correct annual discount rate with which to value the distant future.

Now we outline our derivation of the formula for r_∞. Let us first note that in the continuous limit the discount function defined in Eq. (1) is given by

$$D(t) = \mathrm{E}\left[\exp\left(-\int_0^t r(t')dt'\right)\right], \tag{6}$$

where the expectation $E[\cdot]$ is now an average over all possible interest rate trajectories up to time t. We also observe that in terms of the cumulative process

$$x(t) = \int_0^t r(t')dt' \tag{7}$$

the discount is given by

$$D(t) = E\left[e^{-x(t)}\right].$$

Therefore,

$$D(t) = \int_{-\infty}^{\infty} dr \int_{-\infty}^{\infty} e^{-x} p(x, r, t|r_0) dx, \tag{8}$$

where $p(x, r, t|r_0)$ is the joint probability density function of the bidimensional diffusion process $(x(t), r(t))$. From Eqs. (2) and (7) we see that this bidimensional process is defined by the following pair of stochastic differential equations

$$dx(t) = r(t)dt,$$
$$dr(t) = -\alpha[r(t) - m]dt + kdw(t),$$

which implies that the joint density obeys the following Fokker-Planck equation

$$\frac{\partial p}{\partial t} = -r\frac{\partial p}{\partial x} + \alpha\frac{\partial}{\partial r}[(r-m)p] + \frac{1}{2}k^2\frac{\partial^2 p}{\partial r^2}. \tag{9}$$

Since $x(0) = 0$ and $r(0) = r_0$, the initial condition of this equation is

$$p(x, r, 0|r_0) = \delta(x)\delta(r - r_0). \tag{10}$$

The problem is more conveniently addressed by working with the characteristic function, that is, the Fourier transform of the joint density

$$\tilde{p}(\omega_1, \omega_2, t|r_0) = \int_{-\infty}^{\infty} e^{-i\omega_1 x}dx \int_{-\infty}^{\infty} e^{-i\omega_2 r}p(x, r, t|r_0)dr. \tag{11}$$

Transforming Eqs. (9)-(10) results in the simpler equation:

$$\frac{\partial \tilde{p}}{\partial t} = (\omega_1 - \alpha\omega_2)\frac{\partial \tilde{p}}{\partial \omega_2} - \left(im\omega_2 + \frac{k^2}{2}\omega_2^2\right)\tilde{p}, \tag{12}$$

with

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

The solution of this initial-value problem is given by the Gaussian function

$$\tilde{p}(\omega_1, \omega_2, t) = \exp\Big\{-A(\omega_1, t)\omega_2^2 - B(\omega_1, t)\omega_2 - C(\omega_1, t)\Big\}, \tag{13}$$

where the expressions for $A(\omega_1, t)$, $B(\omega_1, t)$, and $C(\omega_1, t)$ are obtained in the Appendix A.

Once we know the characteristic function $\tilde{p}$ obtaining the discount function is straightforward. In effect, comparing Eqs. (8) and (11) we see that

$$D(t) = \tilde{p}\big(\omega_1 = -i, \omega_2 = 0, t\big). \tag{14}$$

In our case $D(t) = \exp\{-C(-i, t)\}$ which, after using the expression for $C(\omega_1, t)$ given in the Appendix A, finally results in

$$\begin{aligned} \ln D(t) &= -\frac{r_0}{\alpha}\left(1 - e^{-\alpha t}\right) \\ &+ \frac{k^2}{2\alpha^3}\left[\alpha t - 2\left(1 - e^{-\alpha t}\right) + \frac{1}{2}\left(1 - e^{-2\alpha t}\right)\right] - m\left[t - \frac{1}{\alpha}\left(1 - e^{-\alpha t}\right)\right]. \end{aligned} \tag{15}$$

Note hat the exponential terms in Eq. (15) are negligible for large times and as $t \to \infty$ we finally get

$$\ln D(t) \simeq -\left(m - k^2/2\alpha^2\right)t, \tag{16}$$

which is Eq. (4).

B. *The log normal process*

The log normal process is the continuous limit of a generalization of the geometric random walk described in the previous section. It can be written as

$$\frac{dr}{r} = \alpha + k dw(t), \tag{17}$$

where r is the rate, α and k are constants, α may be positive or negative while k is always positive and $w(t)$ is a Wiener process [4]. Eq. (17) can be integrated at once yielding

$$r(t) = r_0 \exp\left\{(\alpha - k^2/2)t + kw(t)\right\}, \tag{18}$$

showing that $r(t)$ is never negative ($r_0 > 0$).

For the log normal model it is not possible to write an exact expression for the discount in real time. It is, however, possible to obtain an analytical and exact expression for the Laplace transform:

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

After lengthy calculations briefly outlined in the Appendix B one obtains

$$\hat{D}(\sigma) = \frac{2\Gamma(\beta(\sigma))}{k^2\Gamma(2\beta(\sigma)+\lambda)} \int_0^\infty \frac{e^{-(2r_0\zeta/k^2+1/\zeta)}}{\zeta^{\beta(\sigma)+\lambda}} F(\beta(\sigma), 2\beta(\sigma)+\lambda, \zeta^{-1}) d\zeta, \tag{19}$$

where $F(a, b, x)$ is a Kummer function and

$$\lambda = 2(1 - \alpha/k^2), \qquad \beta(\sigma) = \frac{1}{2}\Big[1 - \lambda + \sqrt{(1-\lambda)^2 + 4\sigma}\Big].$$

The expression given by Eq. (19) is the farthest we can go from an analytical point of view and the exact analytical inversion yielding $D(t)$ is beyond reach. Nonetheless, the large-time asymptotic expression for $D(t)$ is easily derived, which in any case is the main quantity of interest. Using the Tauberian Theorems –which relate the small σ behavior of $\hat{D}(\sigma)$ with the large t behavior of $D(t)$ (Pitt, 1958)– one finally gets (see Appendix B)

$$D(t) \sim \begin{cases} \text{constant} & \alpha < k^2/2, \\ e^{-\rho t} & \alpha > k^2/2, \\ t^{-1/2} & \alpha = k^2/2. \end{cases} \tag{20}$$

($t \to \infty$). We thus see that when reversion is weaker than fluctuations ($\alpha < k^2/2$) the discount function goes to a constant value as time progresses. However, when reversion is greater than fluctuations ($\alpha > k^2/2$) the discount function has the expected exponential decay ($\rho > 0$, see Appendix C). The critical case $\alpha = k^2/2$ leads to the hyperbolic discount function as obtained by Farmer and Geneakoplos (2009).

[4] In the geometric random walk of Farmer and Geanakoplos (2009) α and k are not independent parameters but rather are constrained so that $\alpha = k^2/2$.

IV. Empirical Estimates

To see how important the uncertainty-persistence effect on long run interest rates is, we collected data for nominal interest rates and inflation for fourteen countries over spans of time ranging from 87 to 318 years, as summarized in Table I, and used these to construct real interest rates. The countries in our sample are: Argentina (ARG, 1864-1960), Australia (AUS, 1861-2012), Chile (CHL, 1925-2012), Germany (DEU, 1820-2012), Denmark (DNK, 1821- 2012), Spain (ESP, 1821-2012), United Kingdom (GBR, 1694-2012), Italy (ITA, 1861-2012), Japan (JPN, 1921-2012), Netherlands (NLD, 1813-2012), Sweden (SWE, 1868-2012), the United States (USA, 1820-2012), and South Africa (ZAF, 1920-2012). Some examples are plotted in Figure 1. Since all but two of our nominal interest rate processes are for ten year government bonds, which pay out over a ten year period, we smooth inflation rates with a ten year moving average, and subtract the annualized inflation index from the annualized nominal rate to compute the real interest rate.

Real rates are nominal rates corrected by inflation. Nominal rates are given by the IG rates (i.e., 10 year Government Bond Yield) except in the cases of Chile and United Kingdom where, due to unavailability, we take the ID rates (i.e., the 10 year Discount rate). We transform the open IG or ID annual rates into logarithmic rates and denote the resulting time series by $b(t)$. Inflation is represented by the Consumer Price Index (CPI) and its log rate is

$$c(t) = \frac{1}{T} \ln \left[\frac{C(t+T)}{C(t)} \right],$$

where $T = 10$ years and $C(t)$ is the time series of the empirical CPI for each country. Finally, the real interest rate, $r(t)$, is defined by

$$r(t) = b(t) - c(t).$$

The recording frequency for each country is either annual or quarterly.

A striking feature observed in many epochs for all countries is that real interest rates frequently become negative, often by substantial amounts and for long periods of time (see Tables II and III). On average, real interest rates are negative one quarter of the time. This makes the log normal real interest rate model less interesting, as well as many other models which assume that interest rates are essentially always positive. Thus we confine our empirical work to the Ornstein-Uhlenbeck model.

We fit the parameters m, k, a of the OU model to each of the data series (see also Appendix C). The resulting parameters are listed in Table II.

The last three columns of Table II give the corresponding long run rate of interest defined in Eq. (5), $r_\infty = m - k^2/2\alpha^2$, along with its maximum and minimum value for each country. One observes that r_∞ is indeed on average 25% lower than m for the nine countries with positive m. For example, in the United States, $m = 2.6\%$ and $r_\infty = 2.1\%$. In the UK, $m = 3.3\%$ and $r_\infty = 2.8\%$. In Argentina, $m = 2.4\%$ and $r_\infty = 1.1\%$.

We explain our method of estimation shortly. Before doing so, we emphasize that this exercise suffers from three problems that are forced on us by compromises of expediency. First, the Ornstein Uhlenbeck model may not be the best model of interest rates to fit the historical data. Indeed our fit shows that the data are not stationary with respect to that

	Country	Consumer Price Index	Bond Yields	from	to	records
1	Italy	CPITAM annual from 12/31/1861 quarterly from 12/31/1919	IGITA10 quarterly	12/31/1861	09/30/2012	565
2	Chile	CPCHLM quarterly	IDCHLM quarterly	03/31/1925	09/30/2012	312
3	Canada	CPCANM quarterly	IGCAN10 quarterly	12/31/1913	09/30/2012	357
4	Germany	CPDEUM annual from 12/31/1820 quarterly from 12/31/1869	IGDEU10 quarterly	12/31/1820	09/30/2012	729
5	Spain	CPESPM annual from 12/31/1821 quarterly from 12/31/1920	IGESP10 quarterly	12/31/1821	09/30/2012	709
6	Argentina	CPARGM annual from 12/31/1864 quarterly from 12/31/1932	IGARGM quarterly	12/31/1864	03/31/1960	342
7	Netherlands	CPNLDM annual	IGNLD10D annual	12/31/1813	12/31/2012	189
8	Japan	CPJPNM quarterly	IGJPN10D quarterly	12/31/1921	12/31/2012	325
9	Australia	CPAUSM annual from 12/31/1861 quarterly 12/31/1991	IGAUS10 quarterly	12/31/1861	09/30/2012	564
10	Denmark	CPDNKM annual from 12/31/1821 quarterly from 12/31/1914	IGDNK10 quarterly	12/31/1821	09/30/2012	725
11	South Africa	CPZAFM quarterly	IGZAF10 quarterly	12/31/1920	09/30/2012	329
12	Sweden	CPSWEM annual	IGSWE10 annual	12/31/1868	09/30/2012	135
13	United Kingdom	CPGBRM annual	IDGBRD* annual	12/31/1694	12/31/2012	309
14	United States	CPUSAM annual	TRUSG10M annual	12/31/1820	10/30/2012	183

Table I List of the data analyzed. Notes (i) Chile: we have taken the Discount (ID) rate since the Government Bond Yield data was not available. (ii) Germany: From 06/30/1915 to 03/31/1916 IGDEU is empty and we have repeated the previous record. (iii) Spain: From 07/31/1936 to 12/31/1940 no records available. 07/31/1936 is empty and we have repeated the previous record. (iv) Netherlands: 2/31/1945 is empty and we have repeated the previous record: (v) Japan: From 12/31/1946 to 09/30/1948 is empty and we have repeated the previous record.

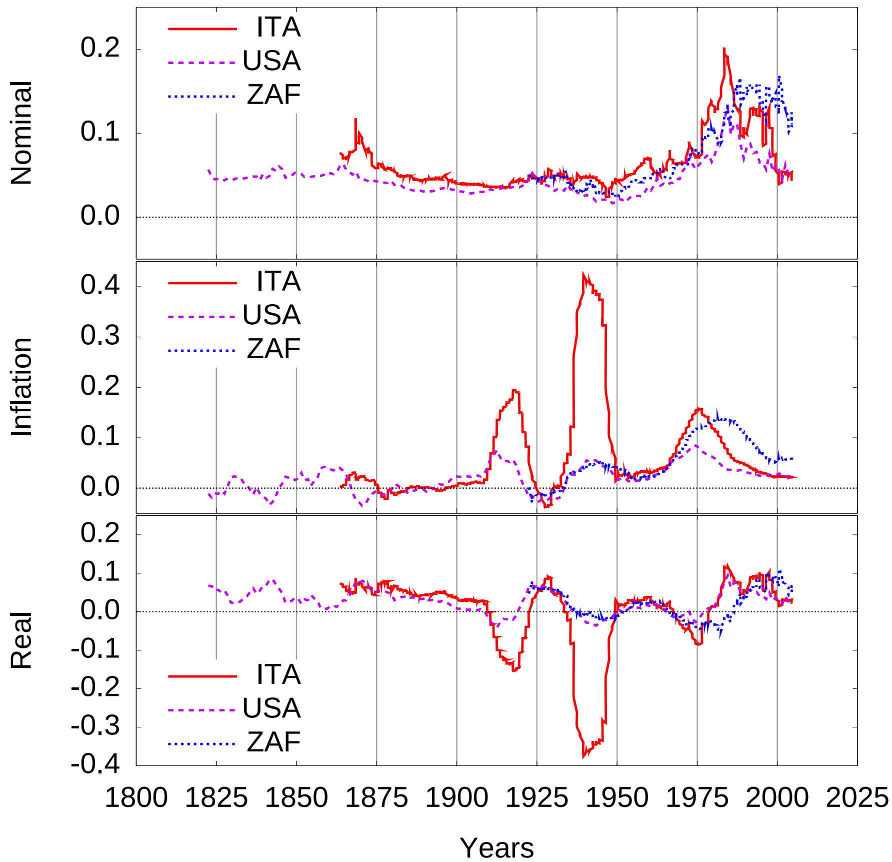

Figure 1. (Color online) Real interest rates display large fluctuations and negative rates are not uncommon. We show nominal interest rates (top), inflation (middle), and real interest rates (bottom) for Italy (ITA), United States (USA) and South Africa (ZAF).

Country	m	Min	Max	k	Min	Max	α	r_∞	Min	Max
Italy	−0.3	−9.1	5.6	6.9	0.8	10.1	0.22(2)	−5.4	−20	5.5
Chile	−6.8	−20.2	12.0	25.2	5.6	44.1	0.40(4)	−26	−74	10
Canada	2.9	0.1	6	2.3	1.1	2.0	0.26(5)	2.5	0.0	5.8
Germany	−10.7	−51.0	4.0	33.9	0.9	61.4	0.20(1)	−160	−540	3.9
Spain	5.7	−0.5	13.5	2.9	1.2	3.6	0.0591(6)	−6.4	−4.8	4.5
Argentina	2.4	−2.9	6.8	6.2	2.8	6.7	0.39(5)	1.1	−4.4	6.5
Netherlands	3.2	0.8	5.4	1.6	0.8	2.2	0.14(1)	2.4	−0.4	5.0
Japan	−2.2	−7.8	4.0	9.7	1.1	13.2	0.24(3)	−10	−23	3.9
Australia	2.6	−0.7	4.9	2.3	0.7	2.8	0.19(2)	1.9	−1.1	4.8
Denmark	3.2	1.5	4.3	2.3	1.1	2.9	0.23(2)	2.7	1.0	4.0
South Africa	1.8	−2.2	5.5	2.5	1.2	2.0	0.21(5)	1.1	−2.3	5.1
Sweden	2.3	−0.3	3.9	2.5	0.6	3.4	0.25(6)	1.9	−0.3	3.8
United Kingdom	3.3	1.4	4.3	1.9	1.0	2.4	0.19(1)	2.8	0.6	4.0
United States	2.6	1.0	4.0	1.8	1.2	2.1	0.18(2)	2.1	0.3	3.8

Table II Parameter estimation of the Ornstein-Uhlenbeck model in yearly units. Notes (i) The columns m, k (in %) and α are estimates taking each country time series; r_∞ (in %) is evaluated from Eq. (5). (ii) The Min and Max columns illustrate the robustness of the estimation procedure by providing the minimum and the maximum value of parameter estimation on four equal length data blocks. (iii) Parenthesis in α column gives the error in the parameter fitting done through linear regression of the autocorrelation function of $r(t)$.

model. Groom et al (2007) find that more complicated models fit better (though of course more complicated models suffer from other problems, including the danger of overfitting). Second, the historical record includes nominal interest rates and inflation, not real interest rates. We derive real interest rates by a crude application of Fisher's equation, subtracting realized inflation from nominal interest rates.[5] Third, the probability distributions assumed in all of the interest rate models we have mentioned, such as the γ_s in Section 2, refer to the so called market probabilities, in which actual probabilities are adjusted by a risk premium determined by the willingness of agents to take bets. Thus $\gamma_{s_u} = \gamma_{s_d} = 1/2$ means that agents are willing to bet on rates going up rather than down at even odds. It does not mean that they think rates are equally likely to go up or down.[6] By using historical frequencies as proxies for the market probabilities in the model we are implicitly assuming that the risk premium does not matter, for example because everybody is risk neutral. Because solving the problems stated above is beyond the scope of the present work, we proceed with the data analysis. Despite their shortcomings, we believe the results nonetheless yield useful insight (Groom et al (2007) and a long literature too numerous to mention find that more complicated models fit better).

We estimate the parameters m, k and α of the OU model as follows: The rate m is the

[5] Freeman et al (2013), among others, pursue an alternative, using cointegration methods to tease out real rates.

[6] There may be many reasons that people are willing to bet at even odds when they think the probabilities are really 2 to 1. For example, such a bet may be a hedge, just like when people buy insurance at actuarially unfair odds because they especially need the money in the contingency the insurance pays.

stationary average of the process (2):

$$\mathrm{E}[r(t)] = m.$$

We estimate α and k based on the autocorrelation function $K(t-t') = \mathrm{E}\left[(r(t)-m)(r(t')-m)\right]$. For the OU process this is (see Appendix C for more details)

$$K(t-t') = \frac{k^2}{2\alpha} e^{-\alpha|t-t'|},$$

and α^{-1} is the correlation time. We estimate α (measured in units of 1/year) by evaluating the empirical auto-correlation and fitting it with an exponential. Once α is determined, the parameter k is obtained from the (empirical) standard deviation, $\sigma^2 = \mathrm{E}\left[(r(t)-m)^2\right]$, which is given by the correlation function since $\sigma^2 = K(0)$. Hence

$$k = \sigma\sqrt{2\alpha}.$$

The countries divide into two very clear groups. Nine countries, with relatively stable real interest rates, have long-run positive rates. Five countries with less stable behavior, in contrast, are in the exponentially increasing region, which implies they have long-run negative rates. (It may not be a coincidence that all five have experienced fascist governments). In four cases the average log interest rate m is negative due to at least one period of runaway inflation; the exception is Spain, which has a (highly positive) mean real interest rate, but still has a long-run negative rate.

In Fig. 2 we show the exact discount function $D(t)$ given by Eq. (15) for all countries as a function of time, illustrating the dramatic difference between the two groups. In most cases the behavior is monotonic; however, it can also be non-monotonic, as illustrated by Argentina, which initially increases and then decreases.

In every case convergence to the long-run rate happens within 30 years, and typically within less than a decade. This is in contrast to other treatments of fluctuating rates, which assume short term rates are always (or nearly always) positive and predict that the decrease in the discounting rate happens over a much longer timescale, which can be measured in hundreds or thousands of years (Newell and Pizer 2003; Weitzman 1998; Gollier et al. 2008; Groom et al. 2007; Farmer and Geanakoplos 2009; Hepburn et al. 2007; Freeman et al. 2013).

It is worth noting how similar the behavior of interest rates is in the nine stable countries. Up to a rescaling of time, the long-run behavior of the model depends only on the two non-dimensional parameters μ and κ, defined as

$$\mu = \frac{m}{\alpha}, \qquad \kappa = \frac{k}{\alpha^{3/2}}. \tag{21}$$

The parameter space can be divided into two regions, as shown in Fig 3. For the region in the upper left, where $\mu > \kappa^2/2$ (or equivalently $m > k^2/2\alpha^2$), the mean interest rate is large in comparison to the noise. The long-run discounting function decays exponentially at rate $r_\infty > 0$. For the region in the lower right $\mu < \kappa^2/2$ and thus $r_\infty < 0$, meaning the discount function $D(t)$ *increases* exponentially. On the boundary, $m = k^2/2\alpha^2$, the long-run interestrate $r_\infty = 0$ and the discount function is asymptotically constant.

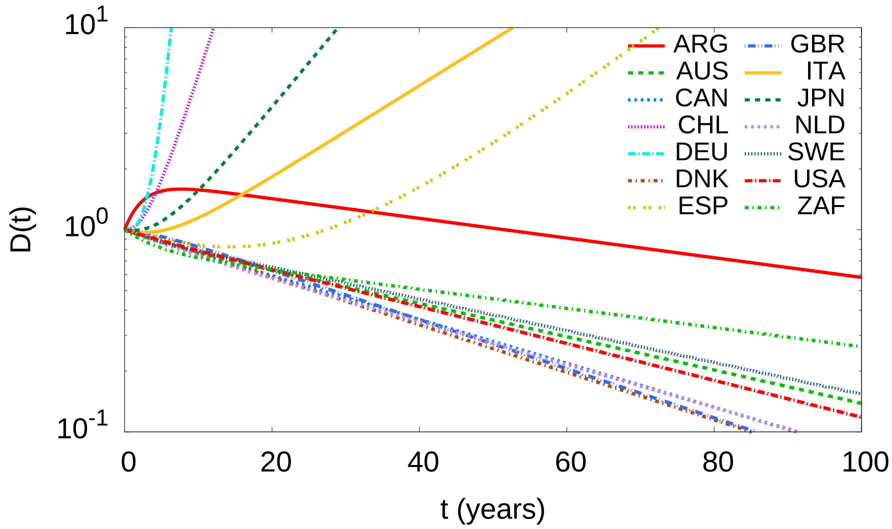

Figure 2. (Color online) The discount function $D(t)$, Eq. (15), as a function of time for the Ornstein-Uhlenbeck model for the fourteen countries in our sample. $D(t)$ quickly reaches its long-run exponential behavior. The long-run rates of the unstable countries vary dramatically, while most of the stable countries are fairly similar.

The position (κ, μ) of each country is shown in Fig 3. The nine countries with long-run positive rates are in the exponentially decaying region at the upper left and are tightly clumped together near the zero long-run interest rate curve. The five countries with long-run negative rates, in contrast, are widely scattered. Note that all fourteen countries are below the identity line in Fig 3, indicating that negative real interest rates are common – even in the stable countries they occur 23% of the time.

This analysis makes it clear why the long-run discount rate is so low. The first reason is that real interest rates are typically fairly low. The average over all countries is 0.71%, and even the average over stable countries (those with $r_\infty > 0$) is 2.7%. The second reason is that the fluctuation term in the second part of Eq. (5), which depends both on the fluctuation amplitude k and the persistence term $1/\alpha$, typically lowers rates for the stable countries by about 22%. In some cases, such as Spain, the effect is much more dramatic: Even though the mean short term rate has the high value of $m = 5.7\%$, the long-term discounting rate is $r_\infty = -6.4\%$. Averaging over the five unstable countries the mean interest rate $m = -2.9\%$ but $r_\infty = -42\%$ (see Table III).

We do not mean to imply that it is realistic to actually use the increasing discounting functions that occur for the five countries with less stable interest rate processes. Hyper-inflation should probably be regarded as an aberration – when it occurs government bonds are widely abandoned in favor of more stable carriers of wealth such as land and gold, and as a result under such circumstances the difference between nominal interest and inflation most likely underestimates the actual real rate of interest.

We mentioned three limitations of our empirical work, including the naive way we deduce real interest rates from nominal interest rates, and the fact we are using historical probabilities instead of market probabilities. Another consideration is that the Ornstein-Uhlenbeck

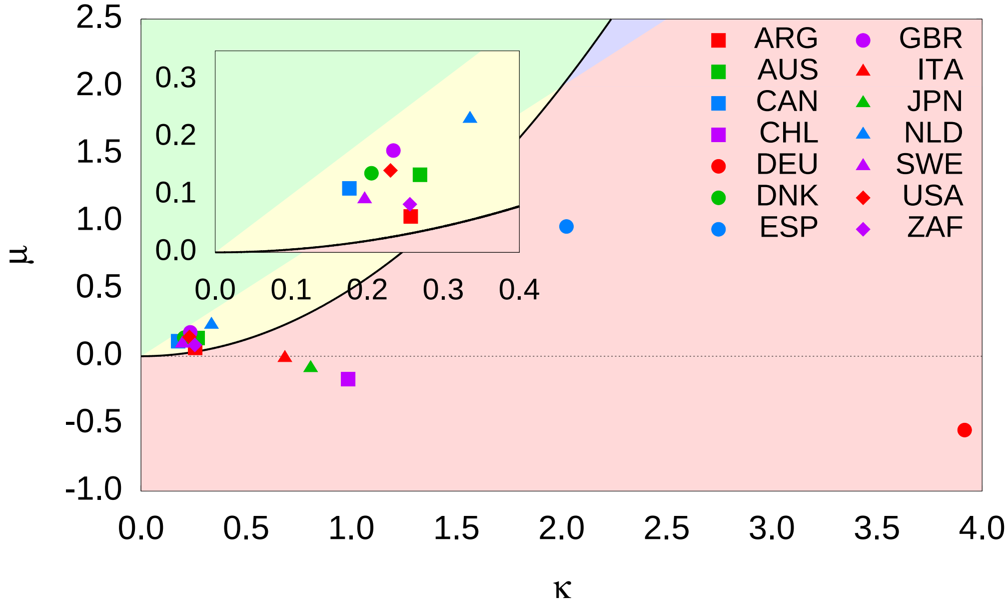

Figure 3. (Color online) A comparison of the parameters of the Ornstein-Uhlenbeck real interest rate model for the fourteen countries in our sample. The vertical axis is the non-dimensional mean interest rate $\mu = m/\alpha$ and the horizontal axis is the non-dimensional fluctuation amplitude $\kappa = k/\alpha^{3/2}$. Points to the upper left of the solid black curve have long-run discount rate $r_\infty > 0$, whereas for those in the lower right $r_\infty < 0$, i.e. the discount function $D(t)$ actually *increases* with time. While the discounting behavior of the nine stable countries is very similar, as shown in the inset, the other five countries behave very differently. Nonetheless, all fourteen countries are below the identity line (the green-yellow interface), indicating that in every case negative real interest rates are common.

Country	Neg RI	$m^{(-)}$%	m%	$1/\alpha$	k	μ	Min	Max	κ	Min	Max	r_∞%	Min	Max
Italy	28% (40y)	13.3	−0.3	4.5	6.9	−0.01	−0.42	0.26	0.68	0.08	1.0	−5.4	−20	5.5
Chile	56% (43y)	25.1	−6.8	2.5	25	−0.17	−0.50	0.30	0.98	0.22	1.7	−26	−74	10
Canada	22% (20y)	1.2	2.9	3.8	2.3	0.11	0.00	0.23	0.18	0.08	0.15	2.5	0.0	5.8
Germany	14% (25y)	100	−10.7	5.0	34	−0.55	−2.6	0.20	3.9	0.10	7.1	−160	−540	3.9
Spain	25% (45y)	3.0	5.7	17	2.9	0.96	−0.08	2.3	2.0	0.85	2.5	−6.4	−4.8	4.5
Argentina	20% (17y)	8.8	2.4	2.6	6.2	0.06	−0.07	0.18	0.26	0.11	0.28	1.1	−4.4	6.5
Netherlands	17% (33y)	1.9	3.2	7.1	1.6	0.23	0.06	0.40	0.34	0.17	0.44	2.4	−0.4	5.0
Japan	33% (26y)	16.1	−2.2	4.2	9.7	−0.09	−0.32	0.17	0.81	0.09	1.1	−10	−23	3.9
Australia	23% (33y)	2.7	2.6	5.3	2.3	0.14	−0.04	0.25	0.27	0.08	0.33	1.9	−1.1	4.8
Denmark	18% (33y)	1.7	3.2	4.3	2.3	0.14	0.07	0.18	0.21	0.10	0.26	2.7	1.0	4.0
South Africa	43% (36y)	0.6	1.8	4.8	2.5	0.08	−0.10	0.26	0.26	0.12	0.21	1.1	−2.3	5.1
Sweden	28% (38y)	1.9	2.3	4.0	2.5	0.09	−0.01	0.15	0.20	0.05	0.27	1.9	−0.3	3.8
U.K.	14% (45y)	0.1	3.3	5.3	1.9	0.18	0.07	0.23	0.23	0.12	0.29	2.8	0.6	4.0
U.S.A	19% (37y)	1.8	2.6	5.6	1.8	0.14	0.05	0.22	0.23	0.16	0.27	2.1	0.3	3.8
All countries	26% (34y)	12.8	0.71	5.4	7.3	0.09	−0.28	0.38	0.75	0.17	1.14	−13.6	−48	5.0
Stable coun.	23% (33y)	2.3	2.7	4.7	2.6	0.13	0.00	0.23	0.24	0.11	0.28	2.1	−0.7	4.8
Unstable coun.	31% (36y)	32	−2.9	6.6	16	0.03	−0.78	0.65	1.67	0.27	2.68	−42	−132	5.6

Table III A summary of our results showing how real interest rates result in a low long-run rate of discounting. This is driven by the fact that average real interest rate m is typically low and the volatility k is substantial. The fact that the characteristic time $1/\alpha$ is typically only a few years implies the long-run discounting rate r_∞ is obtained quickly. Stable countries refer to those with positive r_∞ and unstable countries to those with negative r_∞. Notes (i) "Neg RI" gives the percentage of time and the total number of years in which real interest rates are negative. (ii) $m^{(-)}$ is the average amplitude (in percentage) during negative years only. (iii) m is the mean real interest rate. (iv) $1/\alpha$ is the characteristic reversion time in years. (v) k is the volatility measured in percent. (vi) μ is the non-dimensional mean interest rate. (vii) The Min and Max columns present the minimum and maximum by dividing each series into four equal blocks and estimating parameters separately for each block. (viii) κ is the non-dimensional fluctuation amplitude. (ix) r_∞ is the long-run real interest rate. Negative values of r_∞ mean the discount function is asymptotically increasing.

model, in its simple form, is probably not a good model of interest rates. We confirm this last drawback by testing whether the data is stationary, as assumed by the OU model. In order to have an idea about the robustness of the estimation procedure we split the constructed real interest rate data from each country into four equally spaced blocks (see Appendix C). In each block we estimate the parameters of the OU model applying the method described above, except for the parameter α, which is always estimated using the complete data set. The main reason to avoid estimating α on small blocks is because the time series of some countries are too short. Instead the quoted uncertainty in α is the standard least square error, computed by fitting an exponential to the autocorrelation function of the real interest time series. Tables II and III show the minimum and the maximum values for μ, κ and r_∞, and their uncertainties under subsampling. To provide an estimate of statistical fluctuations we break each country's data into four equal sized blocks and estimate the parameters for each block separately. We quote the maximum and minimum values for each country in Table III. This analysis reveals that statistical uncertainty is large. Focusing on the long-run interest rate r_∞, all countries have positive maximums and most have negative minimums – only the USA, UK, and Denmark have positive r_∞ in all four samples. Subsample variations are more than an order of magnitude larger than standard errors, indicating strong non-stationarity. Our analysis here makes some simplifications, such as ignoring non-stationarity and correlations between the environment and the economy. We believe that including these effects, as we hope to do in future work, will only drive the discounting rate closer to zero. The methods that we have introduced here provide a foundation on which to incorporate more realistic assumptions.

V. Concluding remarks

Financiers have developed a large number of models of interest rate processes to enable them to price bonds and other cash flows. Although these models could in principle be extended to arbitrary horizons, generally they have been studied carefully over a time horizon of about 30 years, since most bonds do not extend much further. Environmental economists, however, are interested in much longer horizons.

The most elementary models of the short run interest rate are the log normal model and the mean reverting normal model of short run interest rates. We derive an analytic expression for the arbitrage free yields on arbitrarily long bonds in these two models by taking their continuous time limits. For the mean reverting normal process –the Ornstein-Uhlenbeck (OU) process– the short run interest rate follows a stationary process with mean m, and variance that by suitable choice of parameters can be taken to be arbitrarily small, maintaining the short rate very close to m with very high probability. Yet we find that with the same parameters the arbitrage free yield on long bonds can be arbitrarily far below m.

We conclude on purely theoretical grounds that economists who wish to do cost benefit analysis over long periods far in excess of 30 years must not hastily use a discount rate set equal to the historical average of the short run rate, even when these rates cluster tightly around their mean. Doing so biases the cost benefit calculations in favor of the present and against interventions that may protect the future.

To get a rough idea of the size of this bias, we look at short run real interest rate data from 14 countries extending over hundreds of years (see Table III). We find that the real interest rate is negative around 20 % of the time, including at the time of the writing of this paper. The fact that the OU model allows for negative real interest rates is therefore a realistic feature. When we fit the parameters of the OU model to the nine countries which never faced a destabilizing hyperinflation, we find an average short rate of about 2.7 % and an average long yield of about 2.1 %.

Let us finish with the following reflection aimed at environmental concerns. Real interest rates are typically closely related to economic growth, and economic downturns are a reality. The great depression lasted for 15 years, and the fall of Rome triggered a depression in western Europe that lasted almost a thousand years. In light of our results here, arguments that we should wait to act on global warming because future economic growth will easily solve the problem should be viewed with some skepticism. When we plan for the future we should always bear in mind that sustained economic downturns may visit us again, as they have in the past.

Appendix A. Discount function for the Ornstein-Uhlenbeck model

We have seen in the main text that when the rate is described by an OU process the joint characteristic function $\tilde{p}(\omega_1, \omega_2, t|r_0)$ obeys the first-order partial differential equation (cf Eq. (12))

$$\frac{\partial \tilde{p}}{\partial t} = (\omega_1 - \alpha\omega_2)\frac{\partial \tilde{p}}{\partial \omega_2} - \left(im\omega_2 + \frac{k^2}{2}\omega_2^2\right)\tilde{p}, \tag{A1}$$

with initial condition

$$\tilde{p}(\omega_1, \omega_2, 0|r_0) = e^{-i\omega_2 r_0}. \tag{A2}$$

We look for a solution of this initial-value problem in the form of a Gaussian density:

$$\tilde{p}(\omega_1, \omega_2, t) = \exp\Big\{-A(\omega_1, t)\omega_2^2 - B(\omega_1, t)\omega_2 - C(\omega_1, t)\Big\}, \tag{A3}$$

where $A(\omega_1, t)$, $B(\omega_1, t)$, and $C(\omega_1, t)$ are unknown functions to be consistently determined. Substituting Eq. (A2) into Eq. (A1), identifying like powers in ω_2 and taking into account Eq. (A2), we find that these functions satisfy the following set of differential equations

$$\dot{A} = -2\alpha A - k^2/2, \qquad A(\omega_1, 0) = 0; \tag{A4}$$

$$\dot{B} = -\alpha B + 2\omega_1 A - im, \qquad B(\omega_1, 0) = ir_0; \tag{A5}$$

$$\dot{C} = \omega_1 B, \qquad C(\omega_1, 0) = 0. \tag{A6}$$

Equation (A4) is a first-order linear differential equation that can be readily solved giving

$$A(\omega_1, t) = \frac{k^2}{4\alpha}\left(1 - e^{-2\alpha t}\right), \tag{A7}$$

Substituting this expression for $A(\omega_1, t)$ into Eq. (A5) results in another first-order equation for $B(\omega_1, t)$, whose solution reads

$$B(\omega_1, t) = ir_0 e^{-\alpha t} + \frac{k^2\omega_1}{2\alpha^2}\Big(1 - 2e^{-\alpha t} + e^{-2\alpha t}\Big) + im\Big(1 - e^{-\alpha t}\Big). \tag{A8}$$

Finally, the direct integration of Eq. (A6) yields the expression for $C(\omega_1, t)$

$$\begin{aligned} C(\omega_1, t) = i\omega_1 r_0 \frac{1}{\alpha}\left(1 - e^{-\alpha t}\right) \;&+\; \frac{k^2\omega_1^2}{2\alpha^3}\left[\alpha t - 2\left(1 - e^{-\alpha t}\right) + \frac{1}{2}\left(1 - e^{-2\alpha t}\right)\right] \\ &+\; im\omega_1\left[t - \frac{1}{\alpha}\left(1 - e^{-\alpha t}\right)\right]. \end{aligned} \tag{A9}$$

From Eq. (14) 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 from Eqs. (A3) and (A9) we obtain

$$\begin{aligned} \ln D(t) = -\frac{r_0}{\alpha}\left(1 - e^{-\alpha t}\right) \;&+\; \frac{k^2}{2\alpha^3}\left[\alpha t - 2\left(1 - e^{-\alpha t}\right) + \frac{1}{2}\left(1 - e^{-2\alpha t}\right)\right] \\ &-\; m\left[t - \frac{1}{\alpha}\left(1 - e^{-\alpha t}\right)\right]. \end{aligned} \tag{A10}$$

Negative rates

As pointed out in the main text, a characteristic of the OU model is the possibility of attaining negative values. This probability is given by

$$P(r > 0, t|r_0) = \int_{-\infty}^{0} p(r, t|r_0) dr, \tag{A11}$$

where $p(r, t|r_0)$ is the probability density function of the rate process. This is given by the marginal density

$$p(r, t|r_0) = \int_{-\infty}^{\infty} p(x, r, t|r_0) dx.$$

The characteristic function of the rate is then related to the characteristic function of the bidimensional process $(x(t), r(t))$ by the simple relation

$$\tilde{p}(\omega_2, t|r_0) = \tilde{p}(\omega_1 = 0, \omega_2, t|r_0).$$

From Eq. (A3) and Eqs. (A7)-(A9) we have

$$\tilde{p}(\omega_2, t|r_0) = \exp\left\{ -\frac{k^2}{4\alpha} \left(1 - e^{-2\alpha t}\right) \omega_2^2 - i \left[r_0 e^{-\alpha t} + m \left(1 - e^{-\alpha t}\right)\right] \omega_2 \right\}.$$

After Fourier inversion we get the Gaussian density[7]

$$p(r, t|r_0) = \frac{(\alpha/k^2)^{1/2}}{\sqrt{\pi(1 - e^{-2\alpha t})}} \exp\left\{ -\frac{(\alpha/k^2)[r - r_0 e^{-\alpha t} - m(1 - e^{-\alpha t})]^2}{1 - e^{-2\alpha t}} \right\}. \tag{A12}$$

The probability for $r(t)$ to be negative, Eq. (A11), is then given by

$$P(r < 0, t|r_0) = \frac{1}{2} \text{Erfc}\left(\frac{(\alpha/k^2)^{1/2}[r_0 e^{-\alpha t} + m(1 - e^{-\alpha t})]}{\sqrt{1 - e^{-2\alpha t}}} \right), \tag{A13}$$

where $\text{Erfc}(z)$ is the complementary error function. Note that as t increases (in fact starting at $t > \alpha^{-1}$) this probability is well approximated by the stationary probability, defined as

$$P_s^{(-)} \equiv \lim_{t \to \infty} P(r < 0, t|r_0).$$

That is

$$P_s^{(-)} = \frac{1}{2} \text{Erfc}\left(m\sqrt{\alpha/k^2} \right). \tag{A14}$$

In terms of the dimensionless parameters μ and κ defined in Eq. (21) this probability reduces to

$$P_s^{(-)} = \frac{1}{2} \text{Erfc}\left(\mu/\kappa \right). \tag{A15}$$

Let us now see the behavior of $P_s^{(-)}$ for the cases (i) $\mu < \kappa$ and (ii) $\mu > \kappa$.

[7] PDF first obtained by G. E. Uhlenbeck and L. S. Ornstein in 1930 (Uhlenbeck and Ornstein, 1930).

(i) If the normal rate μ is smaller than rate's volatility κ, we use the series expansion

$$\mathrm{Erfc}(z) = 1 - \frac{2}{\sqrt{\pi}} z + O(z^2).$$

Hence,

$$P_s^{(-)} = \frac{1}{2} - \frac{1}{\sqrt{\pi}}(\mu/\kappa) + O(\mu^2/\kappa^2). \tag{A16}$$

For μ/κ sufficiently small, this probability approaches 1/2. In other words, rates are positive or negative with almost equal probability. Note that this corresponds to the situation in which noise dominates over the mean. (ii) When fluctuations around the normal level are smaller than the normal level itself, $\kappa < \mu$, we use the asymptotic approximation

$$\mathrm{Erfc}(z) \sim \frac{e^{-z^2}}{\sqrt{\pi} z} \left[1 + O\left(\frac{1}{z^2}\right)\right],$$

and

$$P_s^{(-)} \sim \frac{1}{2\sqrt{\pi}} \left(\frac{\kappa}{\mu}\right) e^{-\mu^2/\kappa^2}. \tag{A17}$$

Therefore, for mild fluctuations around the mean the probability of negative rates is *exponentially small.*

Rates below the long-run rate

It is also interesting to know the probability that real rates $r(t)$ are below the long-run rate r_∞. This is given by

$$P_\infty(t) \equiv \mathrm{Prob}\{r(t) < r_\infty\} = \int_\infty^{r_\infty} p(r, t|r_0) dr.$$

In the stationary regime, $t \to \infty$, we have

$$P_\infty = \int_\infty^{r_\infty} p(r) dr, \tag{A18}$$

where $p(r)$ is the stationary PDF. For the OU model $p(r)$ is obtained from Eq. (A12) after taking the limit $t \to \infty$:

$$p(r) = \frac{1}{\sqrt{\pi}} \left(\frac{\alpha}{k^2}\right)^{1/2} e^{-\alpha(r-m)^2/k^2}. \tag{A19}$$

Substituting Eq. (A19) into Eq. (A18), taking into account the definition of the long-run rate, Eq. (5), and some simple manipulations finally yield

$$P_\infty = \frac{1}{2} \mathrm{Erfc}\left(\frac{k}{2\alpha^{3/2}}\right). \tag{A20}$$

Note that

$$\frac{k}{2\alpha^{3/2}} = \frac{1}{\sqrt{2\alpha}} \sqrt{\frac{k^2}{2\alpha^2}} = \sqrt{\frac{m - r_\infty}{2\alpha}},$$

where we have used the definition (5) Hence

$$P_\infty = \frac{1}{2}\mathrm{Erfc}\left(\sqrt{\frac{m - r_\infty}{2\alpha}}\right) \tag{A21}$$

which gives P_∞ in terms of the ratio between the differential of rates, $m - r_\infty$, and two times the strength of the reversion to the mean. Using the asymptotic estimates of the complementary error function discussed above, we see that this probability is exponentially small if $|m - r_\infty| \to \infty$ with α fixed, or if $\alpha \to 0$ with a fixed differential of rates $|m - r_\infty|$.

Appendix B. Discount function for the log normal process

We will now find the discount function $D(t)$ when the rate $r(t)$ follows the log normal process (17). In this case the dynamics of the bidimensional diffusion process $(x(t), r(t))$ are given by

$$\begin{aligned} dx(t) &= r(t)dt \\ \frac{dr(t)}{r(t)} &= \alpha dt + kdW(t), \end{aligned}$$

where $\alpha > 0$ and $k > 0$ are positive constants. The Fokker-Planck equation for the joint density $p(x, r, t|x_0, r_0)$ of the bidimensional process now reads

$$\frac{\partial p}{\partial t} = -r\frac{\partial p}{\partial x} - \alpha\frac{\partial}{\partial r}(rp) + \frac{1}{2}k^2\frac{\partial^2}{\partial r^2}(r^2 p), \tag{B1}$$

with initial condition

$$p(x, r, 0|x_0, r_0) = \delta(x)\delta(r - r_0). \tag{B2}$$

As before we work with the characteristic function of the bidimensional process $\tilde{p}(\omega_1, \omega_2, t|r_0)$, defined as the Fourier transform of the joint density (see Eq. (11))

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

Fourier transforming Eqs. (B1)-(B2) we get the following partial differential equation for $\tilde{p}$

$$\frac{\partial \tilde{p}}{\partial t} = (\omega_1 + \alpha\omega_2)\frac{\partial \tilde{p}}{\partial \omega_2} + \frac{1}{2}k^2\omega_2^2\frac{\partial^2 \tilde{p}}{\partial \omega_2^2}, \tag{B3}$$

with initial condition

$$\tilde{p}(\omega_1, \omega_2, 0|x_0, r_0) = e^{-\omega_2 r_0}. \tag{B4}$$

In order to proceed further we take the Laplace transform with respect to time t (in addition to the Fourier transform with respect to x and r). We define

$$\begin{aligned} \hat{q}(\omega_1, \omega_2, \sigma|r_0) &= \int_0^\infty e^{-\sigma t}\tilde{p}(\omega_1, \omega_2, t|r_0)dt \\ &= \int_{-\infty}^{\infty} dx e^{-\omega_1 x}\int_{-\infty}^{\infty} dr e^{-\omega_2 r}\int_0^\infty e^{-\sigma t}p(x, r, t|r_0)dt, \end{aligned} \tag{B5}$$

and the initial-value problem (B3)–(B4) collapses into the following ordinary and inhomogeneous differential equation for $\hat{q}$:

$$\omega_2^2 \frac{d^2\hat{q}}{d\omega_2^2} + \left(\frac{2}{k^2}\omega_1 + \alpha\omega_2\right)\frac{d\hat{q}}{d\omega_2} - \frac{2\sigma}{k^2}\hat{q} = -\frac{2}{k^2}e^{-\omega_2 r_0}. \tag{B6}$$

There are boundary conditions implicitly attached to this equation. Indeed, let us note that $\hat{q}(\omega_1, \omega_2 = 0, \sigma|r_0) = \hat{q}(\omega_1, \sigma|r_0)$ corresponds to the Laplace transform of the characteristic function of process $x(t)$ and this distribution exists and is finite. Therefore,

$$\lim_{\omega_2\to 0} \hat{q}(\omega_1, \omega_2, \sigma|r_0) = \text{ finite}. \tag{B7}$$

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

$$\lim_{w_2\to\pm\infty} \hat{q}(\omega_1, \omega_2, \sigma|r_0) = 0. \tag{B8}$$

In order to proceed further we define a new independent variable

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

and a new function

$$\hat{\psi}(\omega_1, \xi, \sigma|r_0) = \xi^{-\beta}\hat{q}(\omega_1, \xi, \sigma|r_0).$$

Then choosing the undefined exponent $\beta = \beta(\sigma)$ as

$$\beta = \frac{1}{2}\left[1 - \lambda + \sqrt{(1-\lambda)^2 + \frac{8\sigma}{k^2}}\right], \tag{B9}$$

we turn Eq. (B6) into an inhomogeneous Kummer equation

$$\xi\frac{d^2\hat{\psi}}{d\xi^2} + (2\beta + \lambda - \xi)\frac{d\hat{\psi}}{d\xi} - \beta\hat{\psi} = -\frac{(2/k^2)}{\xi^{1+\beta}}e^{-2i\omega_1 r_0/k^2\xi}, \tag{B10}$$

where

$$\lambda = 2(1 - \alpha/k^2). \tag{B11}$$

Two independent solutions of the homogeneous Kummer equation corresponding to Eq. (B10) are the confluent hypergeometric functions $F(\beta, 2\beta+\lambda, \xi)$ and $U(\beta, 2\beta+\lambda, \xi)$ (Magnus, Oberhettinger, and Soni, 1966), which allow us to solve the inhomogeneous equation by the method of variation of parameters. The solution obeying the boundary conditions (B7)-(B8) and written in the original variables $\hat{q}$ and ω_2 reads

$$\hat{q}(\omega_1, \omega_2, s|r_0) = \frac{2\Gamma(\beta)}{k^2\Gamma(2\beta+\lambda)}\left(\frac{2\omega_1}{k^2\omega_2}\right)^{\beta}$$
$$\times\left[U\left(\beta, 2\beta+\lambda, \frac{2\omega_1}{k^2\omega_2}\right)\int_0^{\frac{2\omega_1}{k^2\omega_2}} y^{\beta+\lambda-2}e^{-y-2i\omega_1 r_0/k^2 y}F(\beta, 2\beta+\lambda, y)dy\right.$$
$$\left.+F\left(\beta, 2\beta+\lambda, \frac{2\omega_1}{k^2\omega_2}\right)\int_{\frac{2\omega_1}{k^2\omega_2}}^{\infty} y^{\beta+\lambda-2}e^{-y-2i\omega_1 r_0/k^2 y}U(\beta, 2\beta+\lambda, y)dy\right]. \tag{B12}$$

As explained in Sec. 3, the discount function $D(t)$ is obtained by setting $\omega_1 = -i$ and $\omega_2 = 0$ in the characteristic function $\tilde{p}(\omega_1, \omega_2, t|r_0)$. In the present case we know the Laplace transform of the characteristic function, $\hat{q}(\omega_1, \omega_2, s|r_0)$, given by Eq. (B12) and whose analytical inversion yielding $\tilde{p}$ seems to be beyond reach. We therefore obtain the Laplace transform of the discount function,

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

which is given by $\hat{D}(\sigma) = \hat{q}(\omega_1 = -i, \omega_2 = 0, \sigma|r_0)$. Taking the limit $\omega_2 \to 0$ in Eq. (B12) and bearing in mind the following property of Kummer function U (Magnus et al., 1966)

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

we finally obtain

$$\hat{D}(\sigma) = \frac{2\Gamma(\beta)}{k^2\Gamma(2\beta+\lambda)} \int_0^\infty \frac{e^{-(2r_0\zeta/k^2+1/\zeta)}}{\zeta^{\beta+\lambda}} F(\beta, 2\beta+\lambda, \zeta^{-1}) d\zeta. \tag{B13}$$

Asymptotic discount function for the log normal process

We now outline a proof of the asymptotic estimates of $D(t)$ given in Eq. (20) of the main text. In order to find asymptotic expressions of $D(t)$ for large values of time we will use the so-called Tauberian theorems which relate the large time behavior of any function with the small σ of its Laplace transform (Pitt, 1958).

We see from Eq. (B13) that the σ dependence of $\hat{D}(\sigma)$ is through the quantity $\beta = \beta(\sigma)$ defined in Eq. (B9). We, therefore, assume σ small and expand $\beta(\sigma)$ up first order:

$$\beta = \frac{1}{2}\left[(1-\lambda) + |1-\lambda| + \frac{4}{k^2|1-\lambda|}\sigma + O(\sigma^2)\right]. \tag{B14}$$

The asymptotic form $D(t)$ depends on the range of values taken by the parameter λ which, in turn, depends on the ratio α/k^2 (cf Eq. (B11)). We distinguish two regions, $\lambda > 1$ and $\lambda < 1$, separated by the value $\lambda = 1$. In each of these cases discount shows a markedly distinct behavior as time progresses. Note that the case $\lambda > 1$ corresponds to $\alpha < k^2/2$, in other words, the log rate is a supermartingale. For the case $\lambda < 1$ ($\alpha > k^2/2$) the log rate is a submartingale while in the limit case, $\lambda = 1$ ($\alpha = k^2/2$) the log rate is a martingale [8].

(i) *Supermartingale case ($\lambda > 1$)*

Now $\lambda > 1$ and $1 - \lambda = -|1-\lambda|$, hence

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

[8]Indeed, from Eq. (18) we see that the log-rate $\ln[r(t)/r_0]$ has negative expectation if $\alpha < k^2/2$ which means that the log rate is a supermartingale. When $\alpha > k^2/2$ the expectation is positive (submartingale) and $\alpha = k^2/2$ implies a vanishing expectation (martingale).

Recalling the power series definition of Kummer function F

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

and the behavior of the Gamma function $\Gamma(z)$ as $z \to 0$ (Magnus et al., 1966), we see that

$$F(\beta, 2\beta+\lambda, \zeta^{-1}) = 1 + O(\sigma), \quad \zeta^{-\lambda-\beta} = \zeta^{-\lambda}[1+O(\sigma)], \quad \frac{\Gamma(\beta)}{\Gamma(2\beta+\lambda)} = \frac{k^2(\lambda-1)}{2\Gamma(\lambda)}\frac{1}{\sigma} + O(1).$$

Collecting results in Eq. (B13) we get

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

where

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

Recalling the standard limit property of the Laplace transform (Pitt, 1958),

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

we conclude that when $\lambda > 1$ the discount function saturates towards a constant asymptote:

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

(ii) *Subrmartingale case ($\lambda < 1$)*

In this case $|1-\lambda| = 1-\lambda$ and expansion (B14) now reads

$$\beta = 1 - \lambda + \frac{4\sigma}{k^2(1-\lambda)} + O(\sigma^2),$$

and $\lambda + \beta = 1 + O(\sigma)$, hence $\zeta^{-\beta-\lambda} = \zeta^{-1}[1+O(\sigma)]$. Also

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

Expanding the Gamma function terms of Eq. (B13) we have

$$\frac{\Gamma(\beta)}{\Gamma(2\beta+\lambda)} = \frac{k^2/4\gamma}{k^2(1-\lambda)/4\gamma + \sigma} + O(\sigma^2),$$

where $\gamma = 2\psi(2-\lambda) - \psi(1-\lambda)$ and $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the logarithmic derivative of the Gamma function also known as psi, or digamma, function (Magnus et al., 1966). Parameter γ is positive for any $\lambda < 1$. Indeed, from the property $\psi(1+z) = \psi(z) + 1/z$ we see that $\psi(1+z) > \psi(z)$ if $z > 0$ which proves that $\gamma > 0$ for $\lambda < 1$.

Collecting results into Eq. (B13) we have

$$\hat{D}(\sigma) = K_2(r_0)\frac{1}{1+\rho\sigma}[1+O(\sigma)], \tag{B15}$$

where

$$\rho = \frac{k^2(1-\lambda)}{4\gamma} > 0. \tag{B16}$$

and

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

Tauberian theorems (Pitt, 1958) allow us to get the asymptotic behavior of $D(t)$ by means the Laplace inversion of the approximate expression (B15) which, writing it as

$$\hat{D}(\sigma) \sim \frac{K_2(r_0)}{\rho+\sigma} \qquad (\sigma \to 0),$$

can be readily inverted, yielding the exponential decay

$$D(t) \sim K_2(r_0) e^{-\rho t} \qquad (t \to \infty). \tag{B17}$$

(iii) *Martingale case ($\lambda = 1$)*

In this case we don't need to expand $\beta(\sigma)$ in powers of σ because from the definition (B9) we get the simple and exact expression

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

which allows us to write

$$\frac{\Gamma(\beta)}{\Gamma(2\beta+\lambda)} = \frac{k}{\sqrt{2\sigma}}\left[1 + O(\sqrt{\sigma})\right],$$

$$\zeta^{-\lambda-\beta} = \zeta^{-1}[1 + O(\sqrt{\sigma})],$$

and

$$F(\beta, 2\beta+\lambda, \zeta^{-1}) = F(0, 1, \zeta^{-1}) + O(\sqrt{\sigma}) = 1 + O(\sqrt{\sigma}).$$

Collecting results we have

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

where

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

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

$$D(t) \sim \left(\frac{2}{\pi}\right)^{1/2} \frac{K_3(r_0)}{k\sqrt{t}} \qquad (t \to \infty), \tag{B18}$$

which is the hyperbolic discount obtained by Farmer and Geanakoplos (2009).

Appendix C. Parameter estimation and uncertainties

Parameter estimation

Let us recall that the OU model is defined by means of the linear stochastic differential equation

$$dr(t) = -\alpha(r - m) + kdw(t)$$

whose solution is

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

where t_0 is an arbitrary initial time. In what follows we will assume that the process is in the stationary regime. That is to say, we assume the process started in the infinite past (i.e., $t_0 = -\infty$) and all transient effects have faded away. Therefore,

$$r(t) = m + k\int_{-\infty}^{t} e^{-\alpha(t-t')}dw(t'). \tag{C1}$$

The parameter m is easily estimated by noting that since the Wiener process has zero mean the (stationary) mean value of the rate is

$$\mathrm{E}[r(t)] = m. \tag{C2}$$

The estimation of parameters α and k is based on the correlation function, defined by

$$K(t - t') = \mathrm{E}\left[(r(t) - m)(r(t') - m)\right].$$

From Eqs. (C1) and (C2) we write

$$K(t - t') = k^2 e^{-\alpha(t+t')}\int_{-\infty}^{t} e^{\alpha t_1}\int_{-\infty}^{t} e^{\alpha t_2}\mathrm{E}[dw(t_1)dw(t_2)].$$

Taking into account that

$$\mathrm{E}[dw(t_1)dw(t_2)] = \delta(t_1 - t_2)dt_1 dt_2,$$

where $\delta(\cdot)$ is the Dirac delta function, and performing the integration over t_2, we have

$$K(t - t') = k^2 e^{-\alpha(t+t')}\int_{-\infty}^{t} \Theta(t' - t_1)e^{2\alpha t_1}dt_1,$$

where $\Theta(\cdot)$ is the Heaviside step function. In the evaluation the integral we must take into account whether $t > t'$ or $t < t'$. It is a simple matter to see that the final result reads

$$K(t - t') = \frac{k^2}{2\alpha}e^{-\alpha|t-t'|}. \tag{C3}$$

Let us incidentally note that this equation proves that the correlation time of the OU process is given by α^{-1}. Indeed, the correlation time, τ_c, of any stationary random process with correlation function $K(\tau)$ is defined by the time integral of $K(\tau)/K(0)$. In our case

$$\tau_c \equiv \frac{1}{K(0)} \int_0^\infty K(\tau) d\tau = \frac{1}{\alpha}. \tag{C4}$$

Evaluating the empirical auto-correlation from data and fitting it by an exponential (cf. Eq. (C3)) we estimate α (measured in years units) for each country.
The third and last parameter, k, is obtained from the (empirical) standard deviation,

$$\sigma^2 = \mathrm{E}\left[(r(t) - m)^2\right],$$

which is readily given by the correlation function since $\sigma^2 = K(0) = k^2/(2\alpha)$. Hence

$$k = \sigma\sqrt{2\alpha}. \tag{C5}$$

Measuring uncertainties

In order to have an idea about the robustness of the estimation procedure outlined above we have split the real interest rate data from each country into four equally spaced blocks. In each block we have estimated the parameters of the OU model, applying the method described above except for the parameter α, which is always estimated using the complete data set. The main reason to avoid estimating α on each block is because the time series are too short. The quoted uncertainties in α are simply the standard least square error value computed when fitting an exponential the autocorrelation function of the real interest time series. Tables II and III show the minimum and the maximum values, for each parameter and for each country, obtained by subsampling. These can be compared to the value estimated using the whole time series. In the main paper we also present an analogous table with the estimates for μ, κ and r_∞ and their uncertainties, also based on a division into the same four equally spaced blocks.

REFERENCES

Arrow, K. J., M. L. Cropper, C. Gollier, B. Groom, G. M. Heal, R. G. Newell, W. D. Nordhaus, R. S. Pindyck, W. A. Pizer, P. R. Portney, T. Sterner, R. S. J. Tol, and M. L. Weitzman, 2012, How should benefits and costs be discounted in an intergenerational context? the views of an expert panel, Technical report, Resources for the Future, Washington D.C.

Arrow, K. J., M. L. Cropper, C. Gollier, B. Groom, G. M. Heal, R. G. Newell, W. D. Nordhaus, R. S. Pindyck, W. A. Pizer, P. R. Portney, T. Sterner, R. S. J. Tol, and M. L. Weitzman, 2013, Determining benefits and costs for future generations, *Science* 341, 349–350.

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

Dasgupta, P., 2006, *Comments on the Stern Review's Economics of Climate Change* (Cambridge University Press, Cambridge).

Davidson, I., X. Song, and M. Tippett, 2013, Time varying costs of capital and the expected present value of future cash flows, *European Journal of Finance*, (http://dx.doi.org/10.1080/1351847X.2013.802248).

Dybvig, P., J. Ingersoll, and S. Ross, 1996, Long forward and zero-coupon rates can never fall, *Journal of Business* 60, 1–25.

Farmer, J. D., and J. Geanakoplos, 2009, Hyperbolic discount is rational: Valuing the far future with uncertain discount rates, Working paper, SSRN.

Freeman, M. C., B. Groom, E. Panopoulou, and T. Pantellidis, 2013, Declining discount rates and the fisher effect. inflated past, discounted future?, Working paper 129, Center for Climate Change Economics and Policy.

Geanakoplos, J., W. Sudderth, and O. Zeitouini, 2014, Asymptotic behavior of a stochastic discount rate, *Indian Journal of Statistics* 76-A, 150–157.

Gollier, C., P. Koundouri, and T. Pantelidis, 2008, Declining discount rates: Economic justifications and implications for the long-run policy, *Economic Policy* 23, 757–795.

Groom, B., P. Koundouri, P. Panopoulou, and T. Pantelidis, 2007, Discounting distant future: how much selection affect the certainty equivalent rate, *Journal Applied Econometrics* 22, 641–656.

Hepburn, C., P. Koundouri, P. Panopoulou, and T. Pantelidis, 2007, Social discounting under uncertainty: a cross-country comparison, *Journal Environmental Economy and Management* 57, 140–150.

Jouini, E., J. Cvitanic, and M. Musiela, eds., 2001, *Option Pricing, Interest Rates and Risk Management* (Cambridge University Press, Cambridge).

Litterman, R., J. Scheinkman, and L. Weiss, 1991, Volatility and the yield curve, *Journal of Fixed Income* 1, 49–53.

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

Mendelsohn, R. O., 2006, A critique of the stern report, *Regulation* 29, 42–46.

Newell, R., and N. Pizer, 2003, Discounting the distant future: How much do uncertain rates increase valuations?, *Journal Environmental Economy and Management* 46, 52–71.

Nordhaus, W. D., 2007a, Critical assumptions in the stern review on climate change, *Science* 317, 201–202.

Nordhaus, W. D., 2007b, The stern review on the economics of climate change, *Journal Economics Literature* 45, 687–702.

Nordhaus, W. D., 2008, *A Question of Balance* (Yale University Press, New Haven).

Osborne, M. F. M., 1959, Brownian motion in the stock market, *Operation Research* 7, 145–173.

Pitt, H. R., 1958, *Tauberian Theorems* (Oxford University Press, Oxford).

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

Uhlenbeck, G. E., and L. S. Ornstein, 1930, On the theory of brownian motion, *Physical Review* 36, 823–841.

Weitzman, M. L., 1998, Why the far-distant future should be discounted at its lowest possible rate?, *Journal Environmental Economy and Management* 36, 201–208.

Weitzman, M. L., 2007, A review on the stern review on the economics of climate change, *Journal of Economics Literature* 45, 703–724.