Working Paper no. **156**

Discounting the distant future: What do historical bond prices imply about the long term discount rate?

J. Doyne Farmer,
Institute for New Economic Thinking at the Oxford Martin School,
University of Oxford

John Geanakoplos,
Santa Fe Institute, Santa Fe, New Mexico
Department of Economics, Yale University, New Haven CT

Jaume Masoliver,
Departament de Fisica de la Matèria Condensada, Universitat de Barcelona
Universitat de Barcelona Institute of Complex Systems, UBICS

Miquel Montero
Departament de Fisica de la Matèria Condensada, Universitat de Barcelona
Universitat de Barcelona Institute of Complex Systems, UBICS

Josep Perellò
Departament de Fisica de la Matèria Condensada, Universitat de Barcelona
Universitat de Barcelona Institute of Complex Systems, UBICS

Matteo G. Richiardi
Institute for New Economic Thinking at the Oxford Martin School,
University of Oxford
Collegio Carlo Alberto, Moncalieri

December, 2017

Laboratorio R. Revelli, *Collegio Carlo Alberto*
Via Real Collegio, 30 - 10024 Moncalieri (TO)

Tel. +39 011 670.50.60 - Fax +39 011 670.50.61
www.laboratoriorevelli.it - labor@laboratoriorevelli.it

Discounting the distant future: What do historical bond prices imply about the long term discount rate?

J. Doyne Farmer[a,1,1,1,*], John Geanakoplos[e,d], Jaume Masoliver[f,g], Miquel Montero[f,g], Josep Perelló[f,g], Matteo G. Richiardi[a,1,1]

[a]*Institute for New Economic Thinking at the Oxford Martin School, University of Oxford*
[b]*Mathematical Institute, University of Oxford*
[c]*Nuffield College, University of Oxford*
[d]*Santa Fe Institute, Santa Fe, New Mexico*
[e]*Department of Economics, Yale University, New Haven CT*
[f]*Departament de Física de la Matèria Condensada, Universitat de Barcelona*
[g]*Universitat de Barcelona Institute of Complex Systems, UBICS*
[h]*Collegio Carlo Alberto, Moncalieri*

Abstract

We use the theory of bond pricing to study the long term discount rate. Century-long historical records of 3 month bonds, 10 year bonds, and inflation allow us to estimate real interest rates for the UK and the US. Real interest rates are negative about a third of the time and the real yield curves are inverted more than a third of the time, sometimes by substantial amounts. This rules out most of the standard bond pricing models, which are designed for nominal rates that are assumed to be positive. We therefore use the Ornstein-Uhlenbeck model with risk aversion, which allows negative rates and gives a good match to inversions of the yield curve. We derive the discount function using the method of Fourier transforms and fit it to the historical data. The estimated long term discount rate is 1.7% for the UK and 2.2% for the US. The value of 1.4% used by Stern is less than a stan-

*Corresponding author

Email addresses: doyne.farmer@maths.ox.ac.uk (J. Doyne Farmer), john.geanakoplos@yale.edu (John Geanakoplos), jaume.masoliver@ub.edu (Jaume Masoliver), miquel.montero@ub.edu (Miquel Montero), josep.perello@ub.edu (Josep Perelló), matteo.richiardi@spi.ox.ac.uk (Matteo G. Richiardi)

Preprint submitted to Journal of Public Economics *December 22, 2016*

dard deviation from our estimated long run return rate for the UK, and less than two standard deviations of the estimated value for the US. This lends support for substantial immediate spending to combat climate change.

Keywords:
Discounting, environment, interest rates, inflation, Ornstein-Uhlenbeck process

1. Introduction

For environmental problems such as global warming, future costs must be balanced against present costs (Dasgupta, 2004). This is traditionally done by using an exponential discount function with a constant discount rate. The choice of discount rate has generated a major controversy as to the urgency for immediate action. At low discount rates it makes sense to expend resources today to stave off environmental disasters a century in the future. At high discount rates costly action today for the same purpose would appear to be foolish.

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). In an 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%. Stern has been widely criticized for using such a low rate (Nordhaus, 2007b,a; Dasgupta, 2006; Mendelsohn, 2006; Weitzman, 2007; Nordhaus, 2008). This issue surfaced again with the Calderon report in July 2014. What is the right number? And is it even correct to use an exponential function?

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 (Stern, 2014a,b). There are a variety of reasons postulated to cause the need for discounting, including impatience, economic growth, and declining marginal utility. These are embedded in the Ramsey formula (Ramsey, 1928), which forms the basis for a standard approach to discounting the distant future (Arrow et al., 2012). The normative approach helps us understand the underlying reasons for discounting, but it is difficult to make reliable quantitative estimates about the correct rate because the derivations depend on factors that are difficult to measure empirically.

Here we take a positive approach. Assuming that costs and benefits can be reduced to monetary values, the discounting problem is equivalent to bond pricing. A bond is an instrument that one can purchase now that delivers a payment in the future. Similarly, to combat climate change we must spend now in order to receive environmental and economic benefits in the future. If we can quantify both the expenditure required now and the likely cost of inaction in the future, then the price of the corresponding bond gives us an indication of the discount factor.[1]

The interest rate for bonds as a function of their time to maturity is called the *yield curve*. Most bonds have a time to maturity of 30 years or less, but for environmental problems such as climate change we need to know the discount 100 years or more into the future. We don't have data on bonds of such long maturity. Thus we are faced with the problem of inferring the price of long maturity bonds from data on much shorter maturity bonds. Furthermore, the yield curve fluctuates substantially from year to year, so we need sufficient historical time series for reliable statistical inference. In order to do this we need a reasonable model for real interest rates at different maturities.

In addition to the factors that determine the overall level of short term rates, there are two effects influencing long term rates that must be taken into account. The first of these is risk aversion. The far future is less certain than the near future, so all else equal, we expect that longer term bonds bear greater risk, which should imply higher interest rates.

The second effect is more subtle, and is due to the fact that interest rates are uncertain and highly persistent. This effect was originally pointed out in the environmental context[2] by Weitzman (1998) and Gollier et al. (2008). Weitzman and Gollier considered a stylized example in which future real rates

[1]Of course, there are always intangible effects that are difficult to quantify in monetary terms, and one should be suspicious of any procedure that reduces the existence of a species or a human life to a dollar value. But it is nonetheless informative to see what a purely monetary analysis implies.

[2]In fact this effect has been known much longer in the context of bond pricing, see Vasicek (1977). This was also pointed out in a general context by Dybvig et al. (1996).

are unknown today, but starting tomorrow will be fixed forever, i.e. they will be completely persistent, at one of a finite number of values. In this case the long run rate will be dominated by the lowest value, since asymptotically all the other discount factors will be negligible in comparison. To see this concretely, consider two possible rates, r_1 and r_2, with $r_1 > r_2$, and assume they have equal probability. Then the average discount factor at time t in the future is $D(t) = \frac{1}{2}(e^{-r_1 t} + e^{-r_2 t})$. Note that since the sum of two exponentials is not an exponential, the discounting function is no longer an exponential. But for t sufficiently large $D(t) \approx \frac{1}{2}e^{-r_2 t}$, i.e. the discount function becomes approximately exponential with the lower interest rate. This illustrates that when interest rates are uncertain and persistent, lower interest rates tend to dominate. Although this is a stylized example, we will show in a more realistic example how this effect can cause interest rates to decrease with maturity.

This example also illustrates how the presence of heterogeneous rates mean that the discount function is no longer exponential. In fact deviations from exponential behavior can occur even at large times. Farmer and Geanakoplos (2009) used the reflection principle to prove that when the interest rate $r(t)$ follows a geometric random walk, the discount function decays as $K/\sqrt{t}$ for large t, where K is a constant. They called this hyperbolic discounting because the discount factor $D(t)$ obeys the equation of an hyperbola. In the large time limit a hyperbolic function is greater than any exponentially decaying function, showing that there is no positive long run rate of interest in the geometric random walk model. The hyperbola assigns an infinite value to any permanent positive flow of consumption, meaning that the infinite future is infinitely valuable.

Nonetheless, anecdotal evidence suggests that long-term exponential behavior is the typical case. Farmer et al. (2015) examined a variety of different processes, including more general lognormal processes, the Heston process, and the Ornstein-Uhlenbeck process. They found that the case of the simple log-normal process studied by Farmer and Geanakoplos was the only one that did not display long-term exponential behavior. All the other examples deviated from exponential behavior for short times, but eventually converged to an exponential function. This suggests that, while the transient non-exponential behavior can be important for a few decades, the most important question is the long-term discount rate.

To return to the factors that influence the long term rate, both risk aversion and the uncertainty/persistence effect act together. Risk aversion pushes long term rates up while uncertainty/persistence pulls them down, and in general the yield curve is not monotonic. This is born out in the data in the variation of the yield curve from year to year, and we will also show that it can also be true on average.

We are building on earlier empirical work. Litterman et al. (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 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 depends on uncertainty and persistence, depending also on the particular model and parameters. Both Groom et al. and Newell and Pizer calibrated their models against long historical time series of 10 year bonds.

We add to this previous work in several ways. We study century-long records of the prices of 3 month bonds, 10 year bonds, and inflation, for both the USA and the United Kingdom. We find that real interest rates are negative about a quarter of the time, often by substantial amounts. Previous authors have ignored this fact, and instead forced historical interest rates to be positive in order to be consistent with their models.

We instead take the view that negative real interest rates are a fact that cannot be ignored. This leads us to choose the Ornstein-Uhlenbeck (OU) model, which is compatible with negative real interest rates. This can be written

$$dr(t) = -\alpha\big[r(t) - m\big]dt + kdW(t), \tag{1}$$

where $W(t)$ is the standard Wiener process. The parameter k is the volatility and the parameter α determines how fast $r(t)$ reverts to the mean rate m, i.e. the reciprocal $1/\alpha$ is the characteristic reversion time. In the model real interest rates can be arbitrarily negative but they are always pulled back to the rate m.

The formula for the yield curve for the OU model has been derived in the finance literature (see, for instance, Brigo and Mercurio (2006); Duffie (2001); Piazzesi (2009))). We derive this more simply using Fourier transform methods. The previous work in environmental discounting cited above was based on numerical simulations. In contrast, we take advantage of the existence of closed form solutions, which allows us to better estimate the long term discount rate.

When the OU model is used in the finance literature, $r(t)$ generally refers to the nominal interest rate and the mean reversion parameter α is taken to be so strong that $r(t)$ goes negative with probability close to zero.[3] When using nominal rates the existence of negative rates for the OU model is viewed as a disadvantage. But given the fact that real rates are often negative, the logic is reversed. The prevalence of negative real rates makes many of the standard nominal interest rate models inappropriate vehicles for studying real rates, and the OU becomes an obvious choice.

In addition to our treatment of negative rates, we also give an original contribution regarding properly taking risk aversion into account. The previous work described above was calibrated against a single maturity bond (10 years). Risk aversion is a well-accepted notion in bond pricing, and one would expect models that neglect it to underestimate long term interest rates. We fix this by taking risk aversion into account and fit the resulting model to both 3 month and 10 year bond yields, which provide us with two points on the yield curve. Empirically we find that the yield curve is inverted (i.e. the 10 year real interest rate is lower than the 3 month real interest rate) about a third of the time, for both the US. and the UK. This is roughly what our estimated models predict as well.

The net result of our analysis is that we find that, because of the uncertainty/persistence effect, the long range discount rate is close to the real interest rate of ten year bonds. This means that for the UK we find a long term discount rate of 1.7%.

This is very close to the discount rate used by Stern, which is comfortably inside the confidence interval of our estimate. For the U.S. we find the higher rate of 2.2%. For the U.S. our model-consistent error analysis indicates a 5% estimated quantile of 1.5%, which is slightly higher than Stern's value, but still fairly close.

This paper is organized as follows: In Section 2 we present a derivation of the discount function for the OU process. In Section 3 we present our empirical results, and in Section 4 we conclude.

[3]Davidson, Song, and Tippett (2015) examined a square root Ornstein Uhlenbeck model in which $dx(t) = -\alpha x(t)dt + kdW(t)$ 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. (2015) 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 solve for the long rate by studying a partial differential equation using the Feynman-Kac functional, which is quite different from our approach.

2. The process of discounting in continuous time

We now derive the form of the yield curve for the OU model. In the continuous limit the discount function is

$$D(t) = \mathrm{E}\left[\exp\left(-\int_{t_0}^{t} r(t')dt'\right)\right], \tag{2}$$

where t_0 is an arbitrary initial time and the expectation $E[\cdot]$ is an average over all possible instantaneous real rate trajectories up to time t. This is formally identical to the problem of pricing bonds. The price $B(t_0|t_0+t)$ of a zero-coupon bond issued at time t_0 with unit payoff and maturing at time $t_0 + t$ $(t \geq 0)$ is

$$B(t_0|t_0+t) = \mathrm{E}\left[\exp\left(-\int_{t_0}^{t} n(t')dt'\right)\right], \tag{3}$$

where $n(t)$ is the nominal rate. The difference between the two problems is that for discounting we are interested in the real interest rate $r(t)$, whereas for bond pricing we are typically interested in the nominal rate $n(t)$.

2.1. The general framework

Following the same strategy as for bonds, we compute $D(t)$ via a stochastic process model for the time evolution of $r(t)$. The simplest and most common hypothesis consists in assuming that real rates are described by a Markovian process with continuous sample paths. That is, we assume $r(t)$ is a diffusion process whose time evolution is governed by a stochastic differential equation of the form

$$dr(t) = f(r)dt + g(r)dW(t), \tag{4}$$

where $f(r)$ is the drift, $g(r) > 0$ is the noise intensity and $W(t)$ is the standard Wiener process.

In terms of the cumulative process

$$x(t) = \int_{t_0}^{t} r(t')dt', \tag{5}$$

the discount function is given by

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

Therefore

$$D(t) = \int_{-\infty}^{\infty} dr \int_{-\infty}^{\infty} e^{-x} p(x, r, t | x_0, r_0, t_0) dx, \tag{6}$$

where $p(x, r, t|x_0, r_0, t_0)$ is the probability density function (PDF) of the bidimensional diffusion process $(x(t), r(t))$. The measure corresponding to the function p is sometimes referred to as the *data generating measure*. From Eqs. (4) and (5) we see that this bidimensional process is defined by the following pair of stochastic differential equations

$$dx(t) = r(t)dt, \tag{7}$$

$$dr(t) = f(r) + g(r)dW(t), \tag{8}$$

which implies that the joint density obeys the Fokker-Planck equation (FPE)

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

Since $x(t_0) = 0$ and $r(t_0) = r_0$, the initial condition of this equation is

$$p(x, r, t_0|x_0, r_0, t_0) = \delta(x)\delta(r - r_0). \tag{10}$$

Note that $f(r)$ and $g(r)$ do not depend on time explicitly and the process $r(t)$ is time homogeneous. It is therefore invariant under time translations and we can set $t_0 = 0$ without loss of generality.

2.2. The Ornstein-Uhlenbeck process

We now make explicit choices for the functions f and g. For the reasons given in the introduction, we focus on the OU process, which is a standard model from finance that allows negative rates. In finance the OU process was originally proposed by Vasicek (1977) and is sometimes referred to as the Vasicek model. It is a diffusion process with linear drift and constant noise intensity

$$f(r) = -\alpha(r - m), \qquad g(r) = k. \tag{11}$$

The process is thus governed by Equation (1). As we will soon show, the OU process has a stationary normal distribution with mean and standard deviation (m, σ). Letting $r_0 = r(0)$ be the initial return, in the large time limit the probability density function[4] $p(r, t|r_0)$ has mean m and variance $\sigma^2 = k^2/2\alpha$.

2.3. Adding risk aversion

In the context of bond pricing, if investors are risk neutral then prices can reasonably be modeled based on the data generating measure p. This is sometimes called the Local Expectation Hypothesis (LEH) (Cox et al., 1981; Gilles and Leroy, 1986). However a more general assumption is that investors are sensitive to risk, in which case bonds can no longer be priced this way.[5] Following a standard procedure for bond pricing (Vasicek, 1977; Duffie, 2001; Piazzesi, 2009) we take risk into account by adjusting the drift term to account for risk, according to

$$f(r) = -\alpha(r - m) + g(r)q(r), \tag{12}$$

where $q = q(r) \geq 0$ is the *market price of risk*.[6] This raises the effective interest rate, which in the context of bond pricing means that investors are compensated for taking increased risk. The most common assumption consists in taking q constant,[7] in which case the adjusted drift becomes

$$f(r) = -\alpha(r - m^*), \tag{13}$$

[4]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.

[5]Instead they are priced with an artificial probability density funciton, p^*, which is called either the *risk-neutral probability measure* or the *risk-correcting measure*. The two measures p and p^* are related by the market price of risk, which is the extra return per unit of risk that investors demand to bear risk.

[6]The market price of risk may also depend on the current time, $q = q(r, t)$, but here we assume stationarity.

[7]It is also possible to solve the general case in which the market price of risk grows linearly with the rate, i.e. $q(r) = ar + b$, where $a \geq 0$ and b are free parameters, in which case the adjusted drift is $f(r) = -\alpha^*(r - m^*)$, where

$$\alpha^* = \alpha - ka, \qquad m^* = \frac{\alpha m + bk}{\alpha - ka}$$

where

$$m^* = m + \frac{qk}{\alpha}. \tag{14}$$

The result is that the effective mean interest rate m^* is increased relative to the historically observed interest rate m by a constant amount that depends on the volatility, the reversion rate, and the market price of risk.

The solution is found by solving the Fokker-Planck equation (9), which becomes

$$\frac{\partial p}{\partial t} = -r\frac{\partial p}{\partial x} + \alpha\frac{\partial}{\partial r}\big[(r - m^*)p\big] + \frac{1}{2}k^2\frac{\partial^2 p}{\partial r^2}, \tag{15}$$

with the initial condition (10). The problem is more conveniently addressed by working with the characteristic function, that is, the Fourier transform of the joint density

$$\tilde{p}(\omega_1, \omega_2, t|r_0) = \int_{-\infty}^{\infty} e^{-i\omega_1 x}dx \int_{-\infty}^{\infty} e^{-i\omega_2 r}p(x, r, t|r_0)dr. \tag{16}$$

Transforming Eq. (15) results in the simpler equation:

$$\frac{\partial \tilde{p}}{\partial t} = (\omega_1 - \alpha\omega_2)\frac{\partial \tilde{p}}{\partial \omega_2} - \left(im^*\alpha\omega_2 + \frac{1}{2}k^2\omega_2^2\right)\tilde{p}, \tag{17}$$

with (cf. Eq. (10))

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

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

$$\tilde{p}(\omega_1, \omega_2, t) = \exp\Big\{-A(\omega_1, t)\omega_2^2 - B(\omega_1, t)\omega_2 - C(\omega_1, t)\Big\}, \tag{18}$$

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

Once we know the characteristic function $\tilde{p}$ obtaining the discount function is straightforward. Comparing Eqs. (6) and (17) we see that

$$D(t) = \tilde{p}\big(\omega_1 = -i, \omega_2 = 0, t\big). \tag{19}$$

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

We now make a change of notation: For simplicity we have so far assumed that the discount functions is computed at a fixed time $t_0 = 0$ with initial interest rate r_0 for a time t in the future, whereas in what follows we will want to evaluate the discount at different times. Thus we change the notation so that $t_0 \to t$ and $t \to \tau$, and $r_0 \to r(t)$. The discount function is then

$$\begin{aligned} \ln D(\tau) &= -\frac{r(t)}{\alpha}\left(1-e^{-\alpha\tau}\right) - m^*\left[\tau - \frac{1}{\alpha}\left(1-e^{-\alpha\tau}\right)\right] \\ &+ \frac{k^2}{2\alpha^3}\left[\alpha\tau - 2\left(1-e^{-\alpha\tau}\right) + \frac{1}{2}\left(1-e^{-2\alpha\tau}\right)\right]. \end{aligned} \tag{20}$$

When the maturity time τ is small we can approximate $D(\tau)$ by expanding the exponentials to first order using a Taylor series approximation. This yields $\ln D(\tau) \approx r(t)\tau$, as expected. In contrast, in the limit $\tau \to \infty$ the discount function becomes independent of the initial condition. It decays exponentially and can be written in the form

$$D(\tau) \simeq e^{-r_\infty \tau}, \tag{21}$$

where

$$r_\infty = m + \frac{qk}{\alpha} - \frac{k^2}{2\alpha^2}. \tag{22}$$

Thus we see that the long-run discount rate depends on the historical rate m, but this is shifted by two terms. The first shift raises the long-run rate due to the market price of risk. The second shift lowers it by an amount given by the ratio of uncertainty (as measured by k) and persistence (as measured by α). We can trivially rewrite the equation above as

$$r_\infty = m + \frac{k}{\alpha}\left(q - \frac{k}{2\alpha}\right). \tag{23}$$

This makes it clear that whether or not the overall shift in the long-run discount rate is positive or negative depends on the size of the market price of risk in relation to the ratio of the volatility parameter and the reversion rate.

It is not surprising that the market price of risk raises the long term rate, but it is not so obvious that uncertainty and persistence can lower it. Indeed for the OU process it can make it arbitrarily small. 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 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. Of course, this can always be offset by the market price of risk.

It is even possible for the long-run rate to be negative. This is due to the amplification of negative real interest rates $r(t)$. Computation of the discount function involves an average over exponentials, rather than the exponential of an average. As a result, periods where interest rates are negative are 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 substantially reduce the long run interest rate.

To summarize, in the OU model the long-run discounting rate can be much lower than the mean, and indeed can correspond to low interest rates that are rarely observed.

3. Empirical results

We now estimate the OU model for real interest rates on historical data. For this purpose we have collected long historical time series for both short and long run nominal interest rates, as well as inflation, for the United Kingdom and the United States. The properties of the data are summarized in Table 1. For each country we have both three month and ten year interest rates, as well as an inflation index.

3.1. Estimation of real interest rates

We estimate real interest rates as nominal rates corrected by inflation. This is done via an application of Fisher's equation, subtracting realized

Time series	Identifier	frequency	from	to	# records
United Kingdom 3 month Treasuries	ITGBR3D	monthly	12/31/1900	12/31/2012	113
United Kingdom 10 year bonds	IDGBRD	annual	12/31/1694	12/31/2012	309
UK inflation index	CPGBRM	annual	12/31/1694	12/31/2012	309
United States 3 month Treasury bills	ITUSA3D	monthly	01/31/1920	10/30/2012	93
United States 10 year bonds	TRUSG10M	annual	12/31/1820	10/30/2012	183
United States inflation index	CPUSAM	annual	12/31/1820	10/30/2012	183

Table 1: Data sets used in our main analysis.

inflation from nominal interest rates.[8] An alternative way to model real rates by the OU process could have been the modeling of nominal and inflation rates separately; that is, nominal rates by a positive process (for example the Feller process as in the CIR model (Cox et al., 1985; Farmer et al., 2015)) and inflation –which may assume positive and negative values – by the OU process. Such a procedure would need to take into account possible correlations between bond prices and inflation. It is not clear a priori which procedure is better, and the procedure we follow has the virtue of being simpler.

We transform the annual rates into logarithmic rates and denote the resulting time series by $y(t|\tau)$ (with the maturity time τ equal to either 3 month or 10 years). Nominal rates $n(t)$ are then estimated by $n(t) \sim y(t|\tau)$ (see Appendix B for details). The inflation rate $i(t)$ is estimated through the Consumer Price Index (CPI) as

$$i(t) \sim \frac{1}{\tau} \ln \left[\frac{I(t+\tau)}{I(t)} \right], \tag{24}$$

where $I(t)$ is the aggregated inflation up to time t, and $\tau = 10$ years (Appendix B). Finally, the real interest rate $r(t)$ is defined by

$$r(t) = n(t) - i(t). \tag{25}$$

The recording frequency for each country is either annual or quarterly. For ten year government bonds, which pay out over a ten year period, we smooth inflation rates with a ten year forward moving average, and subtract the

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

annualized inflation index from the annualized nominal rate to compute the real interest rate. For the three month bond rates, in contrast, we use the inflation adjustment for the corresponding year (since we do not have inflation adjustments at quarterly frequency). Figure 1 shows nominal rates, inflation and real interest rates for 3 month bonds for the UK and the US, and Figure 2 compares 3 month and 10 year real interest rates.[9]

3.2. Empirical properties of the data

One of the most striking features of these time series is that real interest rates are often negative, in some cases by substantial amounts and for long periods of time. This is evident in the histograms seen in Figure 3. Real 3 month interest rates for the UK are negative 32% of the time, and there are four distinct periods where they drop below or nearly reach −10%. Ten year real interest rates for the UK are negative 38% of the time. For the US real 3 month interest rates are also negative 32% of the time, dropping below −10% during WWII, and 10 year real interest rates are negative 30% of the time. Given that real interest rates are negative about a third of the time, this makes it clear that models such as the log normal process or the CIR model (Cox at al, 1985; Farmer et al 2015) that assume that rates are non-negative are very far from being appropriate. We therefore confine our empirical work to the Ornstein-Uhlenbeck model.

From Figure 3 it is apparent that the distribution of interest rates is heavy tailed. This is particularly true for the short term interest rates. The heavy tails are apparent both because of the excess in the center of the distribution and from the observations in the tail that exceed the normal distribution. Nonetheless, the deviations are not extreme, and the OU process, which has a normal distribution, is at least a reasonable first approximation.

Another striking feature is that the yield curve is often inverted, i.e. the 10 year real interest rate is often lower than the 3 month rate. For the UK the yield curve is inverted slightly more than 50% of the time and for the US it is inverted 32% of the time. Inversions of the yield curve are obviously important for understanding very long term rates. (See Figure 5, where we compare the empirical inversions to those in the OU model).

[9]This procedure assumes that people correctly forecast inflation, i.e. in absence of any knowledge of behavioral bias we are assuming perfect rationality.

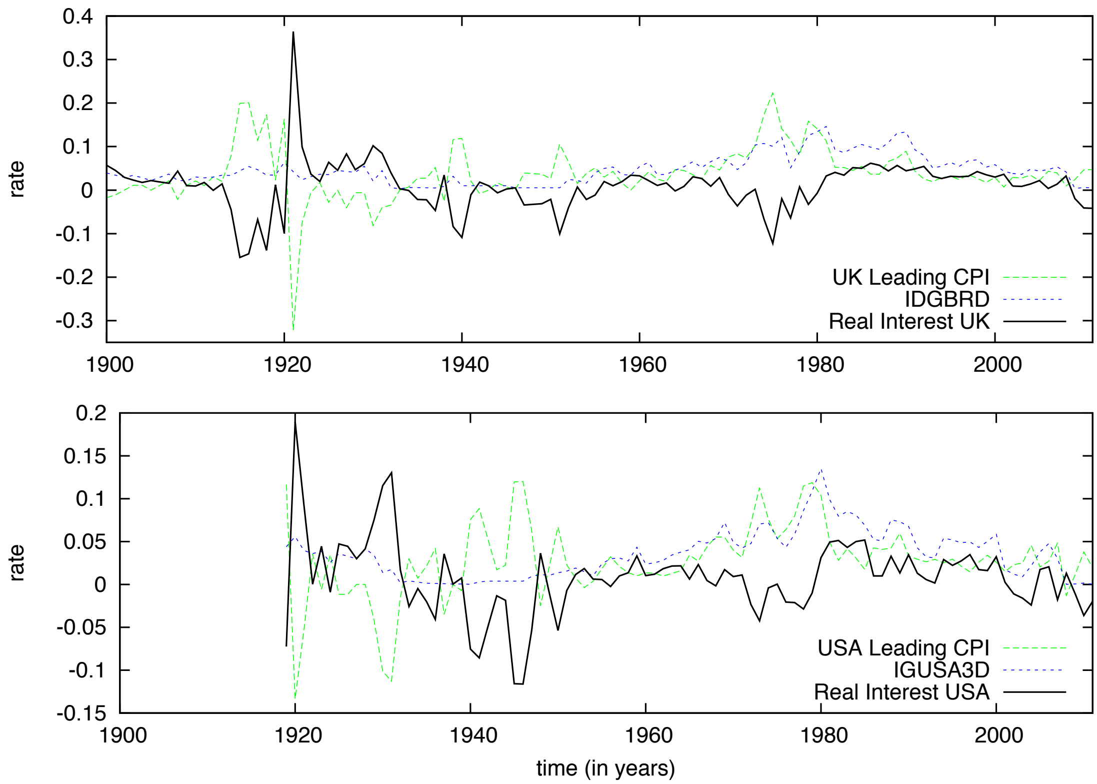
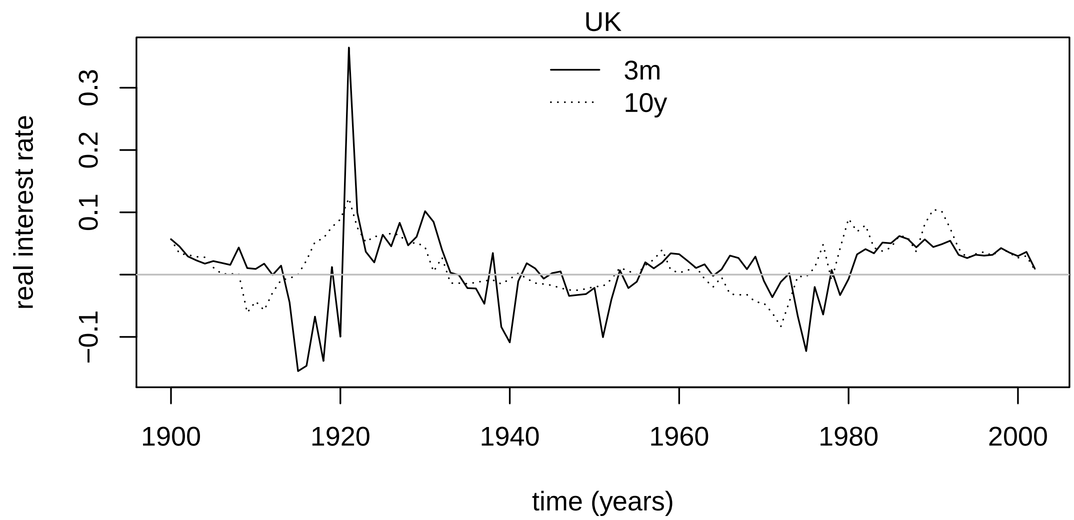


Figure 1: (Color online) Time series for the three month treasury bills for the UK (top) and US (bottom). The inflation index is shown in green (long dashes), the nominal interest rate in blue (short dashes), and the real interest rate is shown as a black solid line. Real interest rates display large fluctuations and negative rates are not uncommon.



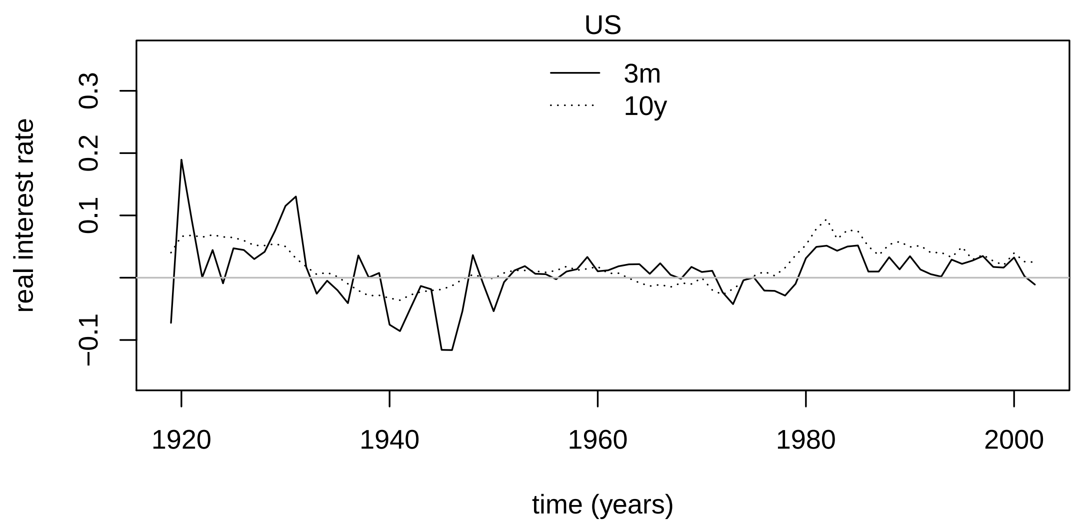


Figure 2: Time series of real interest rates for the UK (top) and the US (bottom). The ten year real interest rates are shown with dashed lines and the three month real interest rates from Figure 1 are shown with solid lines. A horizontal line is drawn at zero to make it clear when real rates are negative.

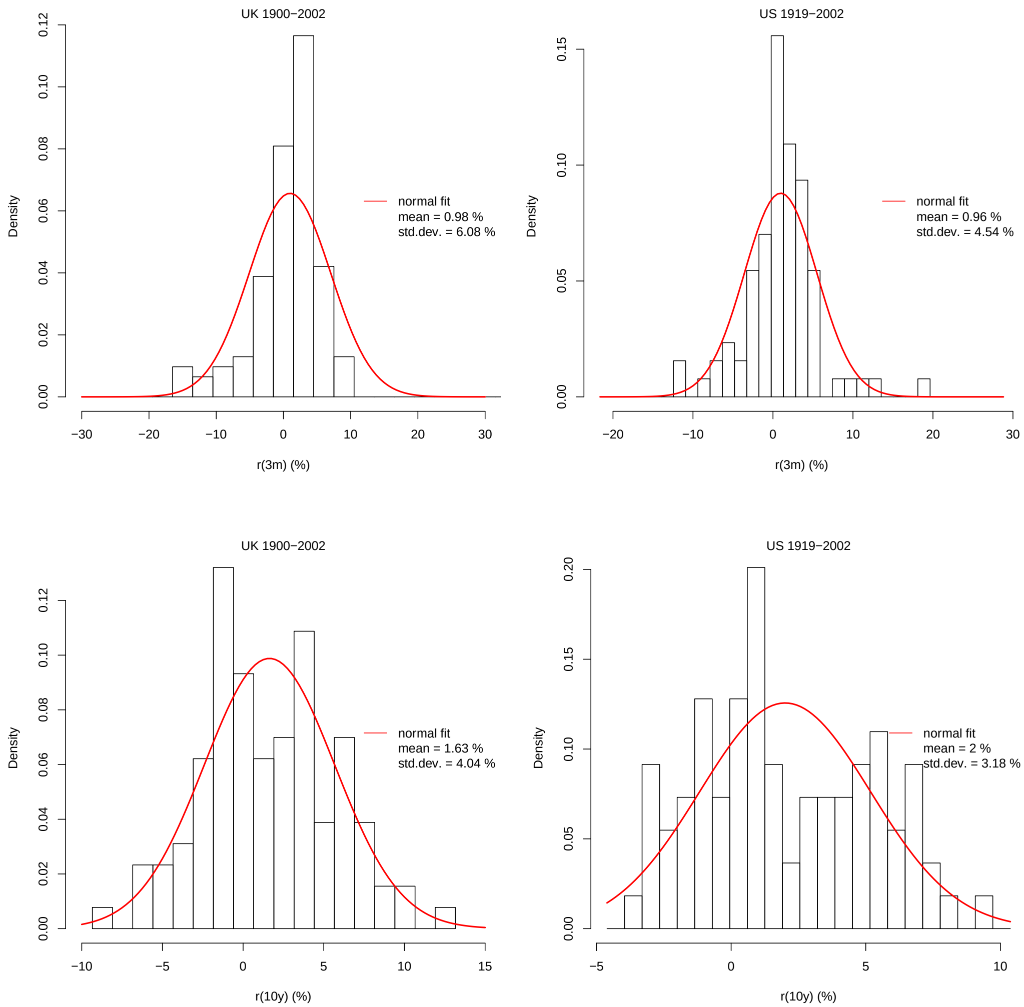


Figure 3: Histograms of 3 month and 10 year real interest rates for the UK and the US. The curve compares to a normal distribution. This illustrates that negative real interest rates are common and that the distribution of interest rates is heavy-tailed relative to a normal distribution.

Country	m	Min	Max	k	Min	Max	α	Min	Max
United Kingdom	0.88	-0.4	3.5	8.2	1.8	15.6	0.93	0.2	1.4
United States	0.83	-1.2	2.2	5.7	2.5	10.5	0.74	0.3	1.3

Table 2: Raw estimates of parameters of the risk-free Ornstein-Uhlenbeck model for the United Kingdom (UK) and the United States (US) based on the annually sampled time series of three-month real interest rates. We use annual units. The Min and Max columns correspond to the minimum and the maximum value of the parameters obtained by splitting the time series into four blocks of equal length and estimating the parameters separately in each block. For better estimates see Table 3.

3.3. Parameter estimation

The OU model with a constant price of risk has four parameters that must be estimated, m, k, α and q. The 3 month rates are much less sensitive to the risk parameter q than the 10 year rates, so by making the approximation that the 3 month is equivalent to the instantaneous process, one can estimate m, k and α from the 3 month rate time series alone. The parameters obtained this way are shown in Table 2.[10]

To provide a feeling for the robustness of the parameter estimation, we divide the time series into four blocks of equal size and estimate the three parameters separately for each block.[11] The maximum and minimum value obtained for each parameter is listed in Table 2. The variations are large, indicating some combination of autocorrelation, non-stationarity and heavy tails. This is evident from the fact that the range of variation in the mean for the four samples is more than three standard deviations.[12]

[10]The values in Table 2 are based on the maximum likelihood estimators derived in Brigo and Mercurio (2006)(see Appendix C).

[11]With the exception of the parameter α , which is always estimated using the complete data set, as the time series in each block are too short. The quoted uncertainties in α are simply the standard least square error value computed when fitting an exponential autocorrelation function of the real interest time series.

[12]If the annual samples were independent, the standard error for measuring the mean would be $SE = \sigma/\sqrt{25}$, where $\sigma = k/\sqrt{2\alpha}$. For the UK this means $SE \approx 1.2\%$. For the four groups of 25 the range of observed values is 3.9%, i.e. a bit more than three standard deviations. For the US (with 20 trials in each sample) $SE = 1.0\%$ and the range is 3.4%,

To measure the q parameter we need to utilize the 10 year real interest rate series as well. The average of the 3 month and 10 year interest rates provide two values that, when used with Eq. (20), gives two independent equations.[13] If we hold k and α fixed based on the previous estimate, we can use these equations to estimate q and improve the estimate of m, obtaining a yield curve that passes precisely through the mean interest rates of the 3 month and 10 year bonds.[14]

The estimates obtained in this manner are slightly distorted because the 3 month bond is only sampled annually, and because we have treated it as though it were an instantaneous rate. We can estimate the size of this bias by simulating the instantaneous process, which we approximate as having daily frequency. We then create a surrogate time series for the 3 month real interest rate time series using Eq. (20) with $\tau = 0.25$ and $r(t)$ as the initial condition at each time t. We then mimic the procedure employed for the real data by estimating the parameters based on the surrogate 3 month rate, sampled at annual frequency. We can then adjust the parameters of the instantaneous process to roughly match, on average, those observed for the real data (i.e. so the estimated values based on the surrogate 3 month series match those of the observed 3 month series). The parameters with the bias corrected are given in Table 3. See Appendix C for more details on the procedure we followed in order to correct the bias. The resulting shift in parameters is small, as can be seen by comparing Tables 2 and 3. The main difference is in the parameter α, which sets the timescale for mean reversion; this changes by a little more than 10%.[15]

---

about three and a half standard deviations.

[13]Putting the average historical 3 month real interest rate and the average 10 year real interest rate on the left hand side of Eq. (20), and substituting $t = 0.25$ and $t = 10$ on the right hand side, gives us two equations. If we take α and k as given, these can be used to estimate the two unknowns m and q. We assume that $r(t) = m$ in Eq. (20), i.e. that the mean historical interest rate is equal to the current rate.

[14]Recall that m is the mean value of the instantaneous process, which is generally different than the average 3 month rate, though we find that the two values are not very different.

[15]Our numerical experiments indicate that the main source of the bias is the annual sampling. The parameter α sets the timescale for mean reversion, so it is not surprising that it is affected by this.

Co.	m	5%	95%	k	5%	95%	α	5%	95%	q	5%	95%
UK	0.84	−0.92	2.6	8.9	7.6	10.4	0.82	0.47	1.26	0.13	−0.04	0.34
USA	0.83	−0.85	2.3	5.8	4.9	6.8	0.65	0.36	1.06	0.20	0.02	0.43

Table 3: Refined estimates of the parameters of the instantaneous Ornstein-Uhlenbeck process using the procedure described in the text. m and k are in percent.

3.4. Comparisons of the model to the data

We now make some comparisons of the simulated OU model model to the real data. Similarly to the procedure for the 3 month rates, we create simulated 10 year rates using Eq. (20) with $r(t)$ as the initial condition at each time t. A comparison shows that the standard deviation of the 3 month rates matches reasonably well, but the standard deviation of the simulated rates is much lower than that of the 10 year rates. We correct for this by adding IID normally distributed noise to make the standard deviation of the simulated and real series match. We also neglect the 10 year smoothing of the 10 year inflation data. The simulated result has a lower correlation between 3 month and 10 year rates than the real data; for the UK the correlations are 21% (simulated) vs. 39% real and for the US 24% (simulated) vs. 59% (real). However the distributions for both 3 month and 10 years match reasonably well. Figure 4 shows the simulated 3 month and 10 year interest rate time series, which should be compared to the real data shown in Figure 2. Not surprisingly, since the 3 month simulated rates are normally distributed, they lack the extreme values observed in the real data.

The OU model does a good job of capturing the frequency of negative interest rates and yield curve inversions. Table 4 compares the frequency of negative interest rates for the real data and the simulation for both 3 month and 10 year rates. In Figure 5 we present a histogram of yield curve inversions for both the data and the model for the UK. (The US histogram is qualitatively similar so we do not present it here). We use the difference between the 10 year real interest rate and the 3 month interest rate as our measure of inversion. The inversions of the data are somewhat more heavy tailed than those of the model, but the agreement is surprisingly good. The real UK yield curve is inverted roughly 50% of the time and the simulated yield curves are inverted 46% of the time. Similarly the real US yield curves is inverted 32% of the time and the simulated yield curves are inverted 41%

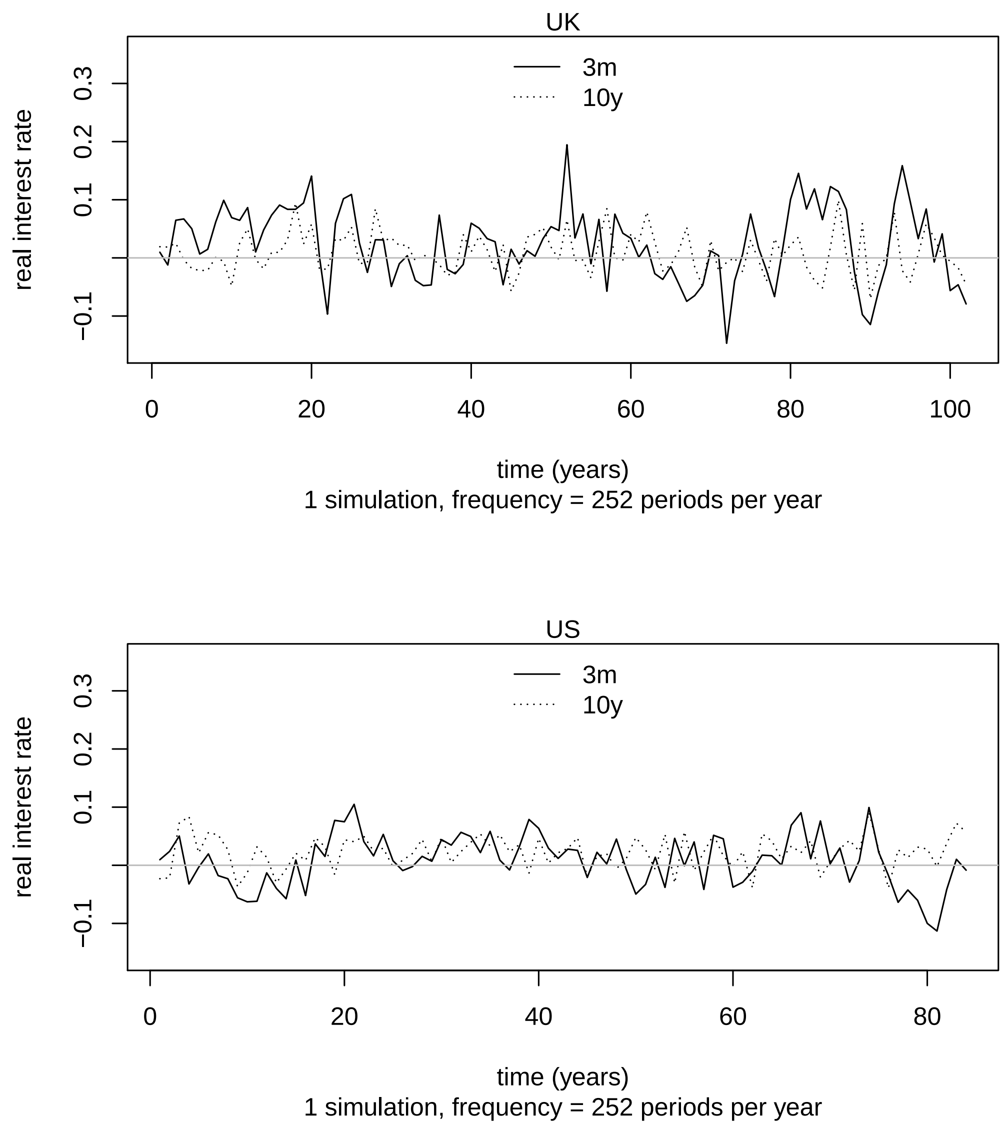


Figure 4: A simulation of the 3 month and 10 year interest rates for the UK (above) and the US (below) using the OU process. Compare to Figure 2.

Country	3 month (data)	3 month (model)	10 year (data)	10 year (model)
UK	32%	43%	38%	34%
USA	32%	42%	30%	26%

Table 4: A comparison of the percentage of the time real interest rates are negative for both the UK and the US in the simulation compared to the data. (The simulation values are averaged over 1000 simulations).

of the time.

3.5. Estimating confidence intervals

With such a short time series it is difficult to estimate confidence intervals by methods such as bootstrapping.[16] However, if we assume that the model is well-specified, then we can at least compute model-consistent error estimates. The width of the resulting confidence intervals can be regarded as lower bounds on the width of the true confidence intervals, and provide a feeling for the magnitude of the estimation errors.

We repeatedly simulate the instantaneous process $r(t)$ using the parameters estimated from the data and generate 3 month and 10 year series as described above. In order to accurately mimic the constraints imposed by the data we sample the simulated series at an annual frequency. We then apply the estimation procedure described above to estimate the four parameters. Doing this 1000 times allows us to compute the 5% and 95% quantiles for each parameter. The results are shown in Table 3.

The estimated discount functions, together with their confidence intervals, are shown in Figure 6. The uncertainty intervals are estimated by repeatedly simulating the instantaneous, 3 month and 10 year processes as described above, applying the estimation procedure to the simulated data, and computing the discount function at each time interval. This is repeated 1000 times to estimate the 5% and 95% quantiles.

We are finally ready to present our key result. The long term interest rate r_∞ is computed using Eq. (22) based on the values in Table 4. The results

[16]This is particularly true for α, where the time series properties of the data matter, so that one would need to do a block bootstrap. There are not very many blocks of sufficient length.

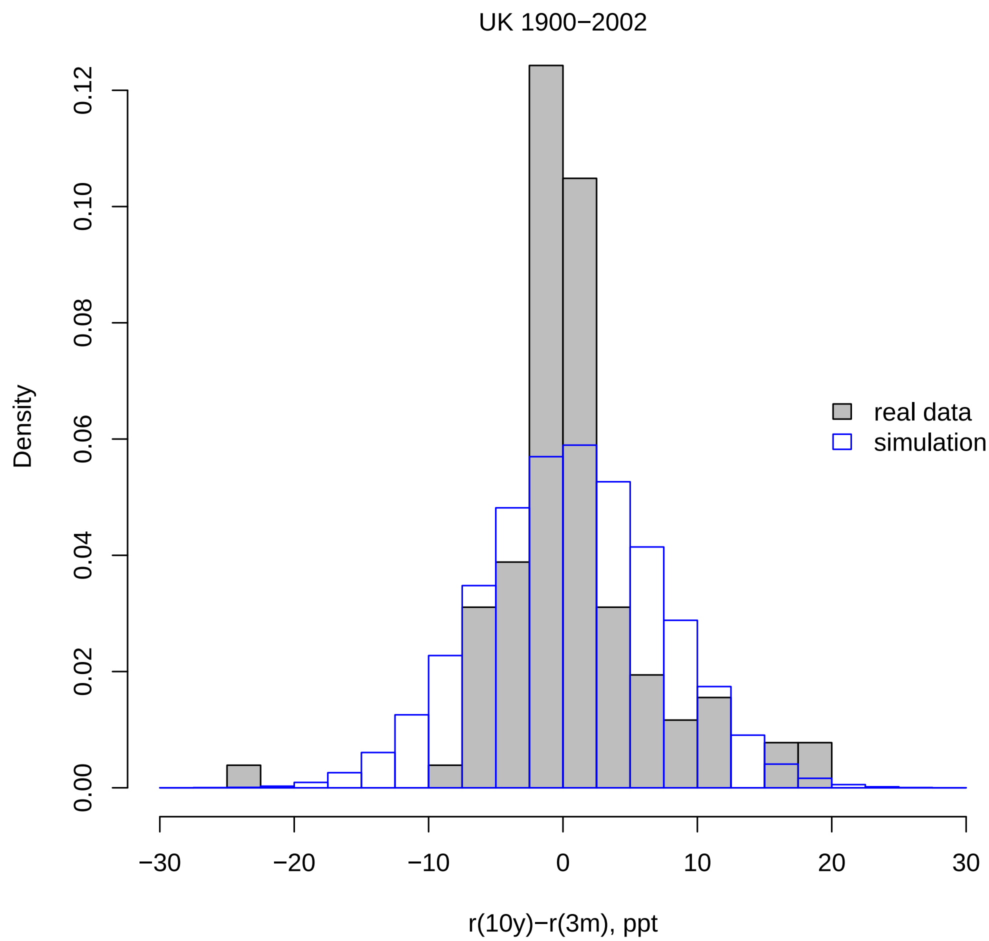


Figure 5: (Color online) Histogram comparing yield curve inversions in the simulated vs. real data for the UK. We measure yield curve inversion based on the difference between the 10 year interest rate and the three month interest rate; positive values indicate a normal yield curve and negative values an inverted yield curve. Interest rates are measured in percent. The real data are shown in grey, the simulation in white. The real data are heavier tailed, but the agreement is otherwise reasonably good.

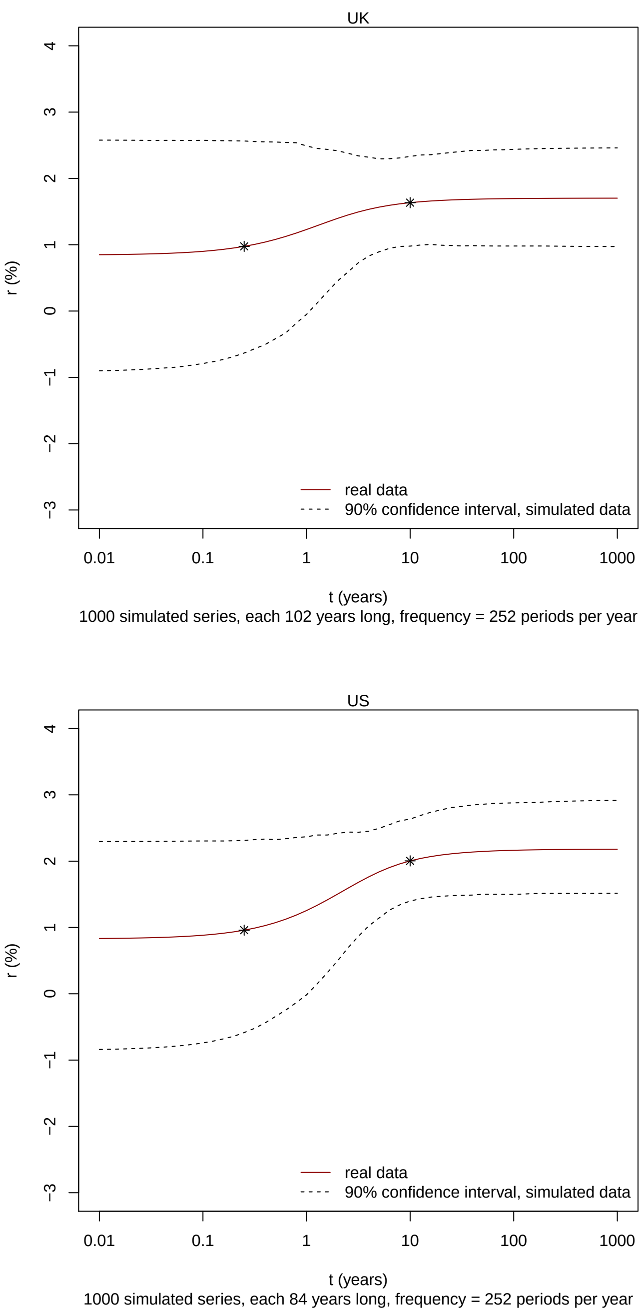


Figure 6: (Color online) The estimated discount functions $D(t)$ for the UK (top) and US (bottom) are shown as solid red lines, plotted versus time (time is on logarithmic scale). The dashed lines indicate the 5% and 95% quantiles based on the simulation procedure described in the text.

Country	r_∞	5%	95%
UK	1.69	0.76	2.63
USA	2.21	1.35	3.07

Table 5: Long term interest rate r_∞ for the United Kingdom and the United States of America, measured in percent, as well as the 5% and 95% quantiles.

are presented in Table 5. The mean long term rate is $r_\infty = 1.7\%$ for the UK and 2.2% for the US. The uncertainties are substantial, with standard deviations of roughly 0.45% in both cases.

4. Concluding remarks

Using historical bond prices to infer long term discount rates is not just a trivial matter of extrapolating mean interest rates, but rather one must take several non-trivial factors into account. To begin with, because real interest rates are so often substantially negative, one must use a model that permits negative rates. This leads us to the Ornstein-Uhlenbeck model. While the presence of negative rates in this model is viewed as a liability for pricing nominal rates, for pricing real rates this becomes a virtue. Another factor that must be taken into account is the market price of risk, which tends to raise longer term rates. Finally one must properly take into account the uncertainty and persistence of interest rates, which tends to lower the long-term discount rate. The use of the OU model accommodates all of these factors. When we estimate the OU model and compare it to the real data, we get a good match for several essential properties, such as the frequencies of negative rates and yield curve inversions.

Our results indicate that the long term interest rate used by Stern is supported by the historical data. His value of 1.4% is less than a standard deviation below the estimated long term rate for the UK of 1.7%, and just under two standard deviations of the US long term rate[17] of 2.2%. In contrast, the interest rate of 4% used by Nordhaus is not supported, as it is well above the 95% confidence intervals of 2.6 for the UK and 3.1 for the US. Our

[17] One should bear in mind that these are the model-consistent confidence intervals; to the extent that the model is mis-specified, these would be widened.

estimates of 1.7% (UK) and 2.2% (US) are compatible with the rates very recently estimated by Giglio, Maggiori and Stroebel (2015). They use data from UK housing markets during 2004–2013 and Singapore during 1995–2013 to estimate an annual discount rate of 2.6 % for payments more than 100 years in the future.[18]

These results could potentially be improved on in several ways. One would be to acquire more data. This could include data for more countries, longer term bonds, or inflation-indexed bonds. Another possible improvement would be to extend the model to better capture the nonstationarity and/or heavy tailed behavior observed in the data.

To conclude, we have demonstrated that historical data indicates that the long-term discount rate is probably not very large. While the error bars remain large, a value of 2% or less seems plausible, corresponding to a present value of about 14% for a payment received 100 years in the future. Values as high as 4% do not appear to be consistent with the historical data.

Acknowledgments

JM, MM, and JP acknowledge partial support from MINECO (Spain) through grants FIS2013-47532-C3-2-P and FIS2016-78904-C3-2-P and also from Generalitat de Catalunya through Complexity Lab Barcelona under contract no. 2014 SGR 608. JDF acknowledges support from the Partners for a New Economy and the Institute for New Economic Thinking.

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. (17))

$$\frac{\partial \tilde{p}}{\partial t} = (\omega_1 - \alpha\omega_2)\frac{\partial \tilde{p}}{\partial \omega_2} - \left(im\omega_2 + \frac{1}{2}k^2\omega_2^2\right)\tilde{p}, \tag{A.1}$$

with initial condition

$$\tilde{p}(\omega_1, \omega_2, 0|r_0) = e^{-i\omega_2 r_0}. \tag{A.2}$$

[18]If we were able to take the observed nonstationarity and/or heavy tails into account, we believe it would decrease the mean values by boosting the uncertainty/persistence effect.

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

$$\tilde{p}(\omega_1, \omega_2, t) = \exp\Big\{-A(\omega_1, t)\omega_2^2 - B(\omega_1, t)\omega_2 - C(\omega_1, t)\Big\}, \tag{A.3}$$

where $A(\omega_1, t)$, $B(\omega_1, t)$, and $C(\omega_1, t)$ are unknown functions to be consistently determined. Substituting Eq. (A.3) into Eq. (A.1), identifying like powers in ω_2 and taking into account Eq. (A.2), we find that these functions satisfy the following set of differential equations

$$\dot{A} = -2\alpha A - k^2/2, \qquad A(\omega_1, 0) = 0; \tag{A.4}$$

$$\dot{B} = -\alpha B + 2\omega_1 A - im\alpha, \qquad B(\omega_1, 0) = ir_0; \tag{A.5}$$

$$\dot{C} = \omega_1 B, \qquad C(\omega_1, 0) = 0. \tag{A.6}$$

Equation (A.4) is a first-order linear differential equation that can be readily solved giving

$$A(\omega_1, t) = \frac{k^2}{4\alpha}\left(1 - e^{-2\alpha t}\right), \tag{A.7}$$

Substituting this expression for $A(\omega_1, t)$ into Eq. (A.5) results in another first-order equation for $B(\omega_1, t)$, whose solution reads

$$B(\omega_1, t) = ir_0 e^{-\alpha t} + \frac{k^2\omega_1}{2\alpha^2}\Big(1 - 2e^{-\alpha t} + e^{-2\alpha t}\Big) + im\Big(1 - e^{-\alpha t}\Big). \tag{A.8}$$

Finally, the direct integration of Eq. (A.6) yields the expression for $C(\omega_1, t)$

$$\begin{aligned} C(\omega_1, t) = i\omega_1 r_0 \frac{1}{\alpha}\left(1 - e^{-\alpha t}\right) \quad &+ \quad \frac{k^2\omega_1^2}{2\alpha^3}\left[\alpha t - 2\left(1 - e^{-\alpha t}\right) + \frac{1}{2}\left(1 - e^{-2\alpha t}\right)\right] \\ &+ \quad im\omega_1\left[t - \frac{1}{\alpha}\left(1 - e^{-\alpha t}\right)\right]. \end{aligned} \tag{A.9}$$

From Eq. (19) 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. (A.3) and (A.9) we obtain

$$\ln D(t) = -\frac{r_0}{\alpha}\left(1-e^{-\alpha t}\right) \quad + \quad \frac{k^2}{2\alpha^3}\left[\alpha t - 2\left(1-e^{-\alpha t}\right) + \frac{1}{2}\left(1-e^{-2\alpha t}\right)\right] \\ - \quad m\left[t - \frac{1}{\alpha}\left(1-e^{-\alpha t}\right)\right]. \tag{A.10}$$

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

$$P(r<0,t|r_0) = \int_{-\infty}^{0} p(r,t|r_0)dr, \tag{A.11}$$

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. (A.3) and Eqs. (A.7)-(A.9) we have

$$\tilde{p}(\omega_2,t|r_0) = \exp\left\{-\frac{k^2}{4\alpha}\left(1-e^{-2\alpha t}\right)\omega_2^2 - i\left[r_0e^{-\alpha t} + m\left(1-e^{-\alpha t}\right)\right]\omega_2\right\}.$$

After Fourier inversion we get the Gaussian density

$$p(r,t|r_0) = \frac{(\alpha/k^2)^{1/2}}{\sqrt{\pi(1-e^{-2\alpha t})}}\exp\left\{-\frac{(\alpha/k^2)[r-r_0e^{-\alpha t}-m(1-e^{-\alpha t})]^2}{1-e^{-2\alpha t}}\right\}. \tag{A.12}$$

The probability for $r(t)$ to be negative, Eq. (A.11), is then given by

$$P(r<0,t|r_0) = \frac{1}{2}\mathrm{Erfc}\left(\frac{(\alpha/k^2)^{1/2}[r_0e^{-\alpha t}+m(1-e^{-\alpha t})]}{\sqrt{1-e^{-2\alpha t}}}\right), \tag{A.13}$$

where $\mathrm{Erfc}(z)$ 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

$$P_s^{(-)} = \frac{1}{2}\mathrm{Erfc}\left(m\sqrt{\alpha/k^2}\right). \tag{A.14}$$

In terms of the dimensionless parameters μ and κ defined by[19]

$$\mu \equiv m/\alpha, \qquad \kappa \equiv k/\alpha^{3/2}, \tag{A.15}$$

this probability reduces to

$$P_s^{(-)} = \frac{1}{2}\mathrm{Erfc}\left(\mu/\kappa\right). \tag{A.16}$$

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

$$P_s^{(-)} = \frac{1}{2} - \frac{1}{\sqrt{\pi}}(\mu/\kappa) + O(\mu^2/\kappa^2). \tag{A.17}$$

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

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

[19]We call μ the dimensionless normal level and κ the dimensionless volatility.

and

$$P_s^{(-)} \sim \frac{1}{2\sqrt{\pi}} \left(\frac{\kappa}{\mu} \right) e^{-\mu^2/\kappa^2}. \tag{A.18}$$

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

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

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

$$P_\infty = \int_\infty^{r_\infty} p(r) dr, \tag{A.19}$$

where $p(r)$ is the stationary PDF. For the OU model $p(r)$ is obtained from Eq. (A.12) after taking the limit $t \to \infty$:

$$p(r) = \frac{1}{\sqrt{\pi}} \left(\frac{\alpha}{k^2} \right)^{1/2} e^{-\alpha(r-m)^2/k^2}. \tag{A.20}$$

Substituting Eq. (A.20) into Eq. (A.19), taking into account the definition of the long-run rate [cf. Eq. (22) with $q = 0$ [20]] and some simple manipulations finally yield

$$P_\infty = \frac{1}{2} \text{Erfc} \left(\frac{k}{2\alpha^{3/2}} \right). \tag{A.21}$$

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 (22). Hence

$$P_\infty = \frac{1}{2} \text{Erfc} \left(\sqrt{\frac{m - r_\infty}{2\alpha}} \right) \tag{A.22}$$

[20]Considering the market price of risk q requires only minor modifications.

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. Empirical estimates. The market price of risk

Recall that real rates are defined as the difference between nominal rates and inflation rates (cf. Eq. (25)):

$$r(t) = n(t) - i(t).$$

We now discuss how to estimate $n(t)$ and $i(t)$ from empirical data .

Nominal rates

Let $B(t|t+\tau)$ be the price at time t of a goverment bond maturing at time $t+\tau$ ($\tau \geq 0$) with unit maturity, $B(t|t) = 1$. The instantaneous rate of return, $b(t|t+\tau)$, of this bond is

$$b(t|t+\tau) \equiv \frac{1}{B(t|t+\tau)} \frac{dB(t|t+\tau)}{dt},$$

so that,

$$B(t|t+\tau) = \exp\left[-\int_t^{t+\tau} b(t|t')dt'\right]. \tag{B.1}$$

It is also useful to define the "yield to maturity", $y(t|\tau)$ as

$$y(t|\tau) \equiv -\frac{1}{\tau} \ln B(t|t+\tau),$$

or, equivalently (cf. Eq. (B.1))

$$y(t|\tau) = \frac{1}{\tau} \int_t^{t+\tau} b(t|t')dt', \tag{B.2}$$

which shows that the yield to maturity is the time average over the maturing period τ of the instantaneous rate of return.

Let us remark that the data at our disposal are not the historical values of $B(t|t+\tau)$ but the annual interest rate of the zero-coupon bond $\beta(t|\tau)$. In this case we have

$$B(t|t+\tau) = \frac{1}{[1+\beta(t|\tau)]^{\tau}}, \tag{B.3}$$

so that

$$y(t|\tau) = \ln[1+\beta(t|\tau)]. \tag{B.4}$$

The spot or nominal rate $n(t)$ is defined as

$$n(t) \equiv \lim_{\tau\to 0} y(t|\tau) = b(t|t), \tag{B.5}$$

where the expression on the right is obtained after substituting for Eq. (B.2) and solving the limit.

Nominal rates are thus estimated by the yield,

$$n(t) \sim y(t|\tau) = \ln[1+\beta(t|\tau)] \tag{B.6}$$

and, attending to definition (B.5), the shorter τ is, the better the estimation for $n(t)$.

Inflation rates

The inflation rate $i(t)$ is estimated by the *ex post* mean inflation rate over a period of time τ, $i(t|\tau)$:

$$i(t|\tau) \equiv \frac{1}{\tau} \ln \frac{I(t+\tau)}{I(t)}, \tag{B.7}$$

where $I(t)$ is the aggregated inflation up to time t. The relation between $I(t)$ and the Consumer Price Index (CPI) is

$$I(t+\tau) = I(t) \prod_{j=0}^{\tau-1} \big[1 + C(t+j)\big], \tag{B.8}$$

where $C(t)$ is the time series of the empirical CPI. The instantaneous rate of inflation $i(t)$ is, therefore, estimated by the quantity $i(t+\tau)$ which written in terms of the CPI reads

$$i(t) \sim i(t+\tau) = \frac{1}{\tau} \sum_{j=0}^{\tau-1} \ln\big[1 + C(t+j)\big]. \tag{B.9}$$

The market price of risk

The concepts of risk neutral probabilities and MPR were developed for bonds and nominal rates. They can be, nonetheless, extended formally to real rates despite practical difficulties because real rates are not tradable and, thus, an empirical basis for constructing a risk neutral measure is lacking.

Let us recall that real rates $r(t)$ are estimated by the quantity $r(t|\tau)$:

$$r(t) \sim r(t|\tau) \equiv y(t|\tau) - i(t|\tau), \tag{B.10}$$

where $y(t|\tau)$ is the yield to maturity τ for a zero-coupon bond $B(t|t+\tau)$ and $i(t|\tau)$ is the inflation rate over period τ.

From theoretical point of view the instantaneous real rate $r(t)$ is defined by

$$r(t) = \lim_{\tau \to 0} r(t|\tau).$$

This leads us to take the shortest possible yield, $y(t|\tau)$, at our disposal ($\tau = 3$ months) to construct a proxy of the real spot rate r.

Obviously the spot rate $r(t)$ is random, so is the quantity $r(t|\tau)$. Let us denote by μ and σ^2 the average and variance of $r(t|\tau)$[21]. Note that in the most general situation $\mu = \mu(t, r|\tau)$ and $\sigma = \sigma(t, r|\tau)$ depend on current time t, rate r, and maturing interval τ (Vasicek, 1977).

The risk premium is defined by the difference $\mu = \mu(t, r|\tau) - r$. Since this excess return depends on the maturity time there can be arbitrage opportunities by buying and selling bonds at different maturities (Vasicek, 1977). It can be shown that these arbitrage opportunities are ruled out as long as the Sharpe ratio of the excess return,

$$q(r,t) \equiv \frac{\mu(t, r|\tau) - r}{\sigma(t, r|\tau)}, \tag{B.11}$$

is independent of the maturity time τ (Vasicek, 1977). This ratio is called

[21] From empirical data these statistics are estimated by

$$\mu \sim \frac{1}{N} \sum_{t=1}^{N} r(t|\tau), \qquad \sigma^2 \sim \frac{1}{N} \sum_{t=1}^{N} \big[r(t|\tau) - \mu \big]^2,$$

where N is the number of samples.

the market price of risk.[22] It depends in general of the current time t and the spot rate r although the most common and feasible assumption is that q is constant or, at most, a function of r.

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

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

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,

$$r(t) = m + k\int_{-\infty}^{t} e^{-\alpha(t-t')}dW(t'). \tag{C.1}$$

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

$$\mathrm{E}[r(t)] = m. \tag{C.2}$$

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

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

[22] Other authors, as for instance Hull et al. (2014), define the MPR as $\lambda = -q$; that is,

$$\lambda = \frac{r - \mu}{\sigma}.$$

From Eqs. (C.1) and (C.2) 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

$$K(t-t') = \frac{k^2}{2\alpha} e^{-\alpha|t-t'|}. \tag{C.3}$$

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

$$\tau_c \equiv \frac{1}{K(0)} \int_0^\infty K(\tau) d\tau = \frac{1}{\alpha}. \tag{C.4}$$

Evaluating the empirical auto-correlation from data and fitting it by an exponential (cf. Eq. (C.3)) 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

$$k = \sigma\sqrt{2\alpha}. \tag{C.5}$$

We estimate these quantities for the three month interest rates using the maximum likelihood procedure given in (Brigo and Mercurio, 2006).

Correcting for the bias from using 3 month rates sampled at annual frequency

Our parameter estimation process that corrects for the bias introduced by using the 3 month rate as an approximation to the instantaneous rate, and by sampling it at annual frequency, has the following steps:

1. Estimate parameters using the historical 3 month and 10 year data as described in the main text.
2. Simulate the instantaneous process (which we approximate as a daily process) using the parameters inferred in step (1) to generate a simulated time series $r(t)$ whose length matches that of the real data (roughly 100 years for the UK and 80 years for the US).
3. Construct simulated 3 month and 10 year time series based on Eq. (20) with $\tau = 0.25$ and $\tau = 10$, using the time series for $r(t)$ from step 1 as the initial condition for each time t.
4. Estimate m, k and α on the simulated 3 month series (sampled at annual frequency).
5. Repeat steps (2-4) for 1000 times and compute the average value of each parameter under the estimation process of step (4). This yields systematic shifts in the parameters relative to those estimated on the historical data, making it clear that the estimation process is biased.
6. Correct for this bias by adjusting the parameters of the instantaneous process by the magnitude of the average shift, so that the estimation process for the simulated 3 month bond times series roughly matches the values estimated from the historical series.

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

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

Brigo, D., Mercurio, F., 2006. Interest Rate Models– Theory and Practice. Springer-Verlag, Berlin and New York.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1981. A re-examination of the traditional hypothesis about the term structure of interest rates. Journal of Finance 36, 769–799.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–407.

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., Song, X., Tippett, M., 2015. Time varying costs of capital and the expected present value of future cash flows. European Journal of Finance 21, 129–146.

Duffie, D., 2001. Dynamic asset pricing theory. Princeton University Press, Princeton.

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

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

Farmer, J. D., Geanakoplos, J., Masoliver, J., Montero, M., Perelló, J., 2015. Value of the future: Discounting in random environments. Physical Review E 91, 052816.

Giglio, S., Maggiori, M., Stroebel, J., 2015. Very long-run discount rates. Quarterly Journal of Economics 130, 1.

Gilles, C., Leroy, S. F., 1986. A note on the local expectation hypothesis. Journal of Finance 41, 975–979.

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

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

Hull, J., Sokol, A., White, A., 2014. Short rates joint measure models. Risk October, 59–63.

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

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

Newell, R., Pizer, N., 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.

Piazzesi, M., 2009. Affine term structure models. In: Sahala, Y. A., Hansen, L. P. (Eds.), The handbook of financial econometrics. Elsevier, pp. 691–766.

Ramsey, F. P., 1928. A mathematical theory of savings. Economic Journal 38, 543–559.

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

Stern, N., 2014a. Ethics, equity and the economics of climate change. paper 1: Science and philosophy. Economics and Philosophy 30, 397.

Stern, N., 2014b. Ethics, equity and the economics of climate change. paper 2: Economics and politics. Economics and Philosophy 30, 445.

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

Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177–188.

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.