Centre for Climate Change Economics and Policy

Grantham Research Institute on Climate Change and the Environment

How certain are we about the certainty-equivalent long term social discount rate?

Mark C. Freeman and Ben Groom

October 2013

Centre for Climate Change Economics and Policy
Working Paper No. 155

Grantham Research Institute on Climate Change and the Environment

Working Paper No. 138

The Grantham Foundation for the Protection of the Environment
Global Green Growth Institute
E·S·R·C Economic & Social Research Council
Munich RE
University of Leeds
LSE The London School of Economics and Political Science

The Centre for Climate Change Economics and Policy (CCCEP) was established by the University of Leeds and the London School of Economics and Political Science in 2008 to advance public and private action on climate change through innovative, rigorous research. The Centre brings together some of the world's leading researchers on climate change economics and policy, from many different disciplines. It was established in 2008 and its first phase ended on 30 September 2013. Its second phase commenced on 1 October 2013. The Centre is funded by the UK Economic and Social Research Council and has five integrated research themes during its second phase:

1. Understanding green growth and climate-compatible development
2. Advancing climate finance and investment
3. Evaluating the performance of climate policies
4. Managing climate risks and uncertainties and strengthening climate services
5. Enabling rapid transitions in mitigation and adaptation

More information about the Centre for Climate Change Economics and Policy can be found at: http://www.cccep.ac.uk.

The Grantham Research Institute on Climate Change and the Environment was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training in climate change and the environment. The Institute is funded by the Grantham Foundation for the Protection of the Environment and the Global Green Growth Institute, and has five research programmes:

1. Global response strategies
2. Green growth
3. Practical aspects of climate policy
4. Adaptation and development
5. Resource security

More information about the Grantham Research Institute on Climate Change and the Environment can be found at: http://www.lse.ac.uk/grantham.

How certain are we about the certainty-equivalent long term social discount rate?

Mark C. Freeman and Ben Groom*

November 4, 2013

Abstract

The case for using declining social discount rates when the future is uncertain is now widely accepted in both academic and policy circles. We present sharp upper and lower bounds for this term structure when we have limited knowledge about the nature of our uncertainty. At horizons beyond 75 years, these bounds are widely spread even if there is agreement on the support and first four moments of the relevant underlying probability distribution. Hence, even in the unlikely event that there is consensus between experts on the primitives of the social discount rate, estimates of the present value of intergenerational costs and benefits, such as the Social Cost of Carbon, can potentially lie anywhere within a wide range. This makes it difficult to prescribe crisp policy recommendations for long-term investments.

JEL Classification: H43, Q51
Keywords: Declining discount rates, Distribution uncertainty, Social Cost of Carbon.

1 Introduction

The outcome of Cost Benefit Analysis (CBA) of public projects with intergenerational consequences is notoriously sensitive to the discount rate employed. Small variations in assumptions can lead to very different policy recommendations. This is particularly problematic because the primatives that underlie discounting analysis are difficult to predict over long

*Mark C. Freeman, School of Business and Economics, Loughborough University, Leicestershire LE11 3TU, United Kingdom. Email: M.C.Freeman@lboro.ac.uk. Ben Groom, Department of Geography and Environment, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom. Email: B.Groom@lse.ac.uk. Author for correspondence: Mark Freeman.

time horizons. For example, the growth rate of aggregate consumption and the rate of return to capital over the next four centuries are essentially unknown today, since they depend on a number of unpredictable events including technological change, political and social unrest, environmental change and even pandemics (e.g. Almond (2006)).

Recognition of the uncertainty surrounding the discount rate has lead to a burgeoning literature on the term structure of certainty-equivalent social discount rates. This has recently been expertly reviewed by Gollier (2012) and Arrow et al. (2012). The overwhelming consensus coming from these contributions is that, for risk free projects, the term structure should be declining with the time horizon. This view is exemplified by a recent *Policy Forum* article in *Science*, in which it is argued that where we are uncertain about the future "there are compelling arguments for using a declining discount rate schedule" (Arrow et al. 2013, p. 350). Furthermore, this consensus has been very influential. Declining term structures can now be found in government guidelines in the UK and France, and in recent advice to the Norwegian government. Declining term structures are also being considered in the US. In the UK, declining discount rate schedules (DDR) have been used in the governmental economic analysis of the High Speed 2 (HS2) rail link and for capital budgeting purposes by the Nuclear Decommissioning Authority. They have also featured heavily in discussions concerning the Social Cost of Carbon (SCC). It can be argued that DDRs have already influenced public investment decisions in the UK.

The DDRs used in government policy are typically based on a 'certainty-equivalent' discount rate which embodies uncertainty in the future. One argument for a declining certainty-equivalent supposes that for some random variable x_H, the present value, p_H, of a certain \$1 arriving at time H is given by $p_H = E\left[\exp\left(-Hx_H\right)\right]$. The $H-$period certainty-equivalent discount rate is then given by $R_H = -H^{-1}\ln\left(p_H\right)$. Exponential functions are convex, and this convexity becomes more pronounced as the time horizon, H, increases and with greater uncertainty in x_H. The certainty-equivalent R_H declines with the time horizon due to the effect of Jensen's inequality to its lowest feasible value as the time horizon extends to infinity, where the limiting value of the certainty-equivalent depends on the persistence of x_H over time. This simple structure, which is the focus of the paper, is quite general since, as we discuss in the next section, x_H can have a number of interpretations depending on the theoretical framework employed.

While this theoretical argument is now widely accepted, putting the theory into practice requires us to make assumptions about the probability density function (pdf) of x_H: $f_H\left(x_H\right)$. The standard approach in the literature is for each author to take one or more specific parameterizations of $f_H\left(x_H\right)$ and apply these as if they reflected the true nature of our knowledge. 'The future is unknown but we can describe our uncertainty pretty accurately' has largely been the approach so far.

We take a less sanguine view of what we can say about the true form of $f_H(x_H)$. We focus on the observation that, not only do we not know what the future holds, but we also cannot perfectly characterize the nature of our uncertainty about it. We most likely have little idea from which distribution the primitives of the discount rate are coming when considering a time-span of many decades or centuries.[1] This leads to our central question: 'how sensitive is the term structure of social discount rates to different plausible characterizations of $f_H(x_H)$?'

While we recognize the philosophy of Knight (1921) that maintains that such uncertainty is immeasurable, we take an alternative approach here.[2] Our framework is based on the willingness of the social planner to make decisions on the basis of specific assumptions about $f_H(x_H)$ that are sufficiently uncontroversial for reasonable people to be able to agree upon them today.[3] Specifically, we consider a set of probability density functions such that all economists agree that $f_H(x_H)$ is a member of this set, yet there is no consensus as to which element of the set best describes $f_H(x_H)$.

A common way in which to describe probability density functions (pdf) is via summary statistics such as the moments of the distribution: e.g. the mean, variance, skewness and kurtosis. Therefore, a reasonable set that people might be able to agree contains $f_H(x_H)$ is that which contains all pdfs defined on the same support that share the same first $K \leq 4$ moments.[4] Of course, the number of pdfs contained in such a set is infinite, but using a powerful method pioneered by Karlin and Studden (1966) it is possible to determine upper and lower limiting pdfs within the set, and hence 'sharp bounds' on the social discount rate based on the level of consensus associated with the set.

Using this technique we present sharp bounds for R_H based on a number of different DDR models. This allows us to then calculate plausible ranges for the present values of three long-term cash flow forecasts; (i) the social cost of carbon, (ii) the estimated benefits of Phase 1 of the HS2 rail link, and (iii) the costs of decommissioning the previous generation of nuclear power stations within the UK.

As is to be expected, the less we agree upon about the future, the more uncertain we are about the "true" present value. For example, using gamma discounting (Weitzman (2001)) as our underlying DDR model, even if there is agreement on the first four moments of the

[1] As we discuss below, in the case of heterogeneous expert opinions, uncertainty over the true form of $f_H(x_H)$ has a somewhat different interpretation.

[2] The derivation of a declining schedule of discount rates in a Knightian uncertainty environment has recently been developed by Iverson (Forthcoming).

[3] Although it is possible that these assumptions will be falsified with the benefit of hindsight.

[4] To avoid issues around infinities, as famously discussed in a related context through the 'dismal theorem' of Weitzman (2009), we assume throughout that the first K moments of $f_H(x_H)$ are finite and that, more generally, its moment generating function is defined.

distribution, then the SCC can lie anywhere within the interval \$13.6 per ton of carbon (/tC) and \$46.1/tC. This is a sobering result when one considers that even agreement on the first moment is likely to be an optimistic assumption. In this position of relative ignorance, the estimate of the SCC can be anywhere between \$5/tC and \$190/tC.

Our conclusion, therefore, is that social planners should be extremely cautious when making policy decisions on intergenerational matters. If we accept that there are elements of the uncertainty surrounding the discount rate that are poorly understood, we can only be sure that the appropriate discount rate will lie within quite wide bounds even if we ignore broader issues such as Knightian uncertainty and the 'dismal theorem' of Weitzman (2009). In such cases spot estimates of, say, the SCC will give a false impression of precision. On the up side our method provides a concrete way in which to set the boundaries of sensitivity analysis in CBA. On the down side, for intergenerational projects we may have to live with the fact that we know very little about the future evolution of the fundamentals and that the social value of such policies may be extremely uncertain. Depressingly for practitioners of CBA, we may have to look elsewhere for a decision making apparatus in these cases.

2 The theory of declining discount rates

In this section, we briefly describe three theoretical foundations for expressions of the form $p_H = E\left[\exp(-Hx_H)\right]$ that result in declining schedules of social discount rates.

2.1 ENPV

We start with the Expected Net Present Value (ENPV) framework of Weitzman (1998) that has spawned the extensive subsequent literature on DDRs. The theoretical foundations for this approach have recently been discussed at length in Traeger (Forthcoming), Gollier and Weitzman (2010), Freeman (2010), Gollier (2009) and elsewhere. Here we followthe discussion in Freeman and Groom (2013).[5] Denote the single period cost of capital, r_t, by the relationship $\exp\left(r_t\right) = E_t[p_{t+1H}]/p_{tH}$, where p_{tH} is the value at time t of a claim to \$1 at time H (so $p_{0H} \equiv p_H$ and $p_{HH} = 1$). Rearranging this to give $p_{tH} = E_t\left[p_{t-1H}\right]\exp\left(-r_t\right)$ and then repeatedly iterating, it is straightforward to derive the initial value $p_0 = E\left[\exp\left(-H\overline{r}_H\right)\right]$, where $\overline{r}_H = H^{-1}\sum_{t=0}^{H-1} r_t$. *Inter alia*, this result is given in (Ang and Liu 2004, equation 17).

[5] This derivation of the ENPV condition significantly relaxes the assumptions of both the resolution of uncertainty and the stochastic process for the interest rate compared to the original thought experiment in Weitzman (1998).

The difficulty within the ENPV framework is placing interpretation on r_t. Following from the literature on the term structure of interest rates this is known to include both a risk-free and a term premium component. However, it is standard in social discounting to treat r_t as a risk-free rate proxied by the yield to maturity on a Treasury bond. We follow this convention in this paper and note that Cox et al. (1981) provide at least some justification for this simplification. Therefore, within the ENPV approach, $x_H = \overline{r}_H$ where $\overline{r}_H$ is the average short-term interest rate over the interval $[0, H-1]$.

2.2 Consumption based asset pricing models

Consider a standard consumption based asset pricing framework where a representative agent gains utility $u(c_t, t)$ from consuming c_t units of the single consumption good at time t. From the Euler equation, if a project makes a certain future payment of \$1 at time H and nothing at any other time, then its present value is given by:

$$p_H = \frac{E\left[u'(c_H, H)\right]}{E\left[u'(c_0, 0)\right]}$$

If we assume that current consumption, c_0, is non-stochastic and that the utility function takes time-separable power form: $u'(c_t, t) = e^{-\rho t} c_t^{-\gamma}$, with pure time preference rate ρ and coefficient of relative risk aversion γ, then:

$$p_H = E\left[e^{-\rho H}\left(\frac{c_H}{c_0}\right)^{-\gamma}\right]$$

In line with our focus in this paper, this can be re-written as:

$$p_H = E\left[\exp\left(-H x_H\right)\right]$$

where $x_H = \rho + \frac{\gamma}{H}\ln\left(c_H/c_0\right)$.

As Gollier (2012) shows, the term structure that emerges from this framework depends on what is assumed about the stochastic process driving consumption growth. Several models are admissible including straightforward mean reversion with a persistent state variable, regime shifting models and models with parameter uncertainty in the mean and variance of the growth process.

2.3 Heterogeneous agent models

While much of the discounting literature focusses on uncertainty, another strand considers heterogeneity, either of preferences or opinions on the discount rate. Doing so leads to analytically similar expressions for the term structure, albeit with an entirely different interpretation.

Suppose there is an economy with $i = 1, .., N$ agents in which each has her own discount rate r_{iH}, for horizon H. In this case, the social planner must decide on the optimal way of combining these different rates. This is a contentious issue and not all approaches lead to expressions of the form $p_H = E[\exp(-Hx_H)]$. For example, Heal (2012) argues that the median value of r_{iH} might be more appropriate due to its democratic median-voter properties. Yet two heterogeneous agent models do lead to such expressions.

in Gollier (2012), equations (9.20) and (9.21) follow Emmerling (2010) and combine the heterogeneous agents by taking a weighted average of their individual discount factors:

$$p_H = \sum_{i=1}^{N} \widehat{q}_i \exp(-Hr_{iH}) \qquad (1)$$

$$\widehat{q}_i = q_i \frac{u'(c_{i0}, 0)}{\sum_{j=1}^{N} q_j u'(c_{j0}, 0)}$$

where the weights q_i are the Pareto weights placed on agent i's utility function by the social planner. Once again, this takes a form appropriate for the focus of this paper: $p_H = E[\exp(-Hx_H)]$, where the expectation is taken with respect to the probability measure $\widehat{q}_i$ and $x_H = r_{iH}$.

Second, following Gollier and Zeckhauser (2005), Jouini et al. (2010) imagine a setting where each agent has his or her own rate of pure time preference, $\rho_i > 0$, and forecast of future economic growth, g_i, but all have logarithmic utility. They then determine their individual discount rate using the Ramsey rule: $r_{iH} = \rho_i + g_i$. In this case, if the social discount rate is determined by equilibrium market prices that would result from each agent trading their own initial endowment, w_i, according to their individual preferences, then:

$$p_H = \sum_{i=1}^{N} z_i \exp(-Hr_{iH}) \qquad (2)$$

$$z_i = \frac{w_i \rho_i}{\sum_{j=1}^{n} w_j \rho_j}$$

Again, setting $x_H = r_{iH}$, this has same structure: $p_H = E[\exp(-Hx_H)]$ where the expectation is now taken with respect to the probability measure z_i. Li and Löfgren (2000) and Heal and Millner (2013) look at alternative specifications of the intertemporal welfare function under heterogeneity and end up with similar expressions.

Thinking about heterogeneity presents perhaps less intractable problems with regard to the future, since in principle one could undertake a survey of the relevant preferences or beliefs. Nevertheless, it would be extremely unrealistic to say that we can either characterize $\widehat{q}_i$ or z_i with precision, leading again to uncertainty about the true form of $f_H(x_H)$.

3 Method

In order to construct term structures of the social discount rate for use in policy decision making, several attempts have been made to characterize $f_H(x_H)$ for use in the equation $p_H = E[\exp(-x_H H)]$. We will turn to these empirical exercises in the following section. Yet despite the best efforts of these endeavours, the precision presented in the eventual social discount rate schedules obscures the difficulty of accurately reflecting our uncertainty about the future and the nature of heterogeneity.

But what is it reasonable to say that we 'know' or agree on about $f_H(x_H)$? The method that we now describe starts from a position of ignorance. We then gradually formalize several scenarios which vary in what is 'known' (in the sense of what reasonable people might agree upon for policy making purposes) about the moments and support of this pdf. We then explain how the extent of our ignorance, as measured by the number of moments of the pdf that are 'known', affects the precision with which we can estimate the shape and level of the term structure of certainty equivalent social discount rates.

First consider the set of all well-defined probability density functions, $\Im_H$, with elements g_H, which are supported on a common interval $[a_H, b_H]$.[6] We assume that there is consensus that the "true" $f_H(x_H)$ is an element of this set, but we do not know which element it is.

Next, we suppose that we only know the first K (non-central) moments of about $f_H(x_H)$; $E_f\left[x_H^k\right] = m_{kH}$ for $k \leq K$ where $E_f[\cdot]$ is the expectation operator conditional on the pdf of x_H being f_H. The smaller is K, the more ignorant we are about $f_H(x_H)$, but claiming any knowledge of the moments of the pdf allows us to narrow our search to the subset $\Im_{KH} \subset \Im_H$ that contains all elements g_H with first K moments $E_g\left[x_H^k\right] = m_{kH}$ for $k \leq K$. We then

[6] There is a restriction that a_H is finite. For all the examples we consider we also take finite b_H, but the extension to infinite b_H is straightforward given the results in Eckberg (1977).

define strict upper and lower bounds for R_H, R_{uH} and R_{lH}, by:

$$\begin{aligned} R_{uH} &= -\frac{1}{H}\ln\left(\inf\left[E_g\left[\exp(-Hx_H)\right]|g_H\in\Im_{KH}\right]\right) \\ R_{lH} &= -\frac{1}{H}\ln\left(\sup\left[E_g\left[\exp(-Hx_H)\right]|g_H\in\Im_{KH}\right]\right) \end{aligned} \tag{3}$$

As expressions of the form $E\left[\exp\left(-Hx_H\right)\right]$ are moment generating functions (mgf), or Laplace-Stieltjes transformations, we can invoke a powerful result from Karlin and Studden (1966) to find R_{uH} and R_{lH}. These are derived by establishing two separate, discrete pdfs which are, loosely speaking, 'at opposite ends' of the support and yet share the first K moments. The upper bound for the mgf is found by calculating the most extreme discrete distribution to place as much mass as possible in the left hand tail (lower values of the discount rate), while still satisfying the K moment conditions. The lower bound is found by minimizing the mass in the left hand tail.

More concretely, the extreme discrete distributions place non-zero probability mass at ϖ points on the interval $[a_H, b_H]$ where the number of mass points depends on the number of moments, K, of the distribution of x_H that we are willing to assume we agree upon: $\varpi\in[(K+1)/2,\ (K+3)/2]$. We denote these points by V_{qlH} (v_{qlH}) for the lower bound of the mgf and V_{quH} (v_{quH}) for the upper bound when $a_H=0$ $(a_H\neq0)$, with associated probabilities π_{qlH} and π_{quH}, where q indexes the mass points from smaller to larger values of x_H. In this way it is possible to define 'sharp bounds' on the term structure of the discount rate based on different levels of ignorance about the future.

Consider the restricted case when $a_H=0$ and $b_H=B_H$; we broaden the discussion to more general values of a_H and b_H in the appendix. The method of Karlin and Studden (1966) now follows from the observation that the set of functions $\left\{1, x_H, ..., x_H^K, (-1)^{K+1}\exp\left(-Hx_H\right)\right\}$ for $H>0$ and positive integer K is a Tchebycheff system. This allows us to identify the properties of the discrete distributions that give sharp bounds for the mgf (see, for example, Eckberg (1977)):

<table>
<tr><th></th><th>V_{qlH}</th><th>V_{quH}</th></tr>
<tr><td>K even</td><td>$\varpi=(K+2)/2$
$K/2$ points in $(0,B_H)$
One point at B_H</td><td>$\varpi=(K+2)/2$
$K/2$ points in $(0,B_H)$
One point at 0</td></tr>
<tr><td>K odd</td><td>$\varpi=(K+1)/2$
$(K+1)/2$ points in $(0,B_H)$</td><td>$\varpi=(K+3)/2$
$(K-1)/2$ points in $(0,B_H)$
One point at 0
One point at B_H</td></tr>
</table>

This is sufficient to uniquely identify each extreme discrete distribution since the number of degrees of freedom equals the number of moment constraints. For example, for even K, there are $K/2$ degrees of freedom on location and $(K+2)/2-1$ degrees of freedom for the probabilities, giving a total number of degrees of freedom of K. It is also straightforward to verify that there are K degrees of freedom in parameter choice when K is odd. Therefore the extreme density functions are uniquely defined by the K moment conditions in all cases.

To illustrate this technique with an example, suppose that experts agree on the first two moments of the distribution; $K = 2$. In this case the extreme distribution that generates the lower bound for $E[\exp(-Hx_H)]$ contains finite mass at the upper bound of $x_H : V_{2lH} = B_H$. There are now two unknowns; the value V_{1lH}, and its associated probability mass, π_{1lH}. V_{1lH} must be chosen from the interval $(0, B_H)$. Given the mean and variance restrictions this point must lie below the mean, and its placement and the probability assigned must satisfy the agreed first two moments for the set of distributions. The upper bound extreme distribution places more emphasis on the left hand tail of the distribution, with one of the mass points at zero and the other mass point and probability chosen in relation to this.

The appendix describes the closed form solutions for V_{qlH}, V_{quH}, π_{qlH} and π_{quH}, which characterize the extreme discrete distributions for the cases $K = 1$, $K = 2$ and $K = 3$; the latter two are found in Eckberg (1977). As noted by (Johnson and Taaffe 1993, p.96), "less analytically tractable cases (e.g., four or five non-central moments) call for use of symbolic or numerical methods for solving the nonlinear equations". We use numerical methods here.[7]

Clearly, the divergence between the sharp bounds on the moment generating function is greater if we claim only to know the first moment of $f_H(x_H)$. If we are willing to make stronger claims on the characteristics of $f_H(x_H)$ and suppose that we know two or more moments of the distribution, the bounds become narrower. The less ignorant we are about the future, the more certain we can be about the term structure of social discount rates.

Of course it is not our contention that economists will perfectly agree on the first K moments of $f_H(x_H)$, nor even its likely support. So, while we present results that vary in the extent of agreement over the moments of the distribution, even the scenario in which we make the least stringent claim to know only the first moment of $f_H(x_H)$ may exhibit a fanciful level of knowledge, particularly for distant time horizons. However, as we now show, even if we claim to know more than is realistic, the potential range over which the social discount rate might lie, $[R_{lH}, R_{uH}]$ for any far horizon H, is often very wide. Introducing further disagreement/ignorance amongst experts will only make this range wider.

[7] Details available on request from the authors.

4 Calibration

In order to calculate the sharp bounds on the social discount rate, several ingredients are needed. Firstly, we must calculate the support of x_H for each time horizon H; a_H and b_H. Second, we must determine the non-central moments of the distribution at each time horizon, $m_{kH} = E\left[x_H^k\right]$, in order to evaluate the impact of varying degrees of ignorance about the future. We limit our analysis to $K \in [1,4]$ as discussions about probability density functions rarely go beyond kurtosis.

The approach we take employs six examples from the literature. These cover the three theoretical frameworks described in Section 2 and all result in a single declining term structure of the social discount rate. For all horizons $H \leq 400$ years, we estimate a_H, b_H and m_{kH} from the baseline probability density function, $F_H(x_H)$, that was used in each case; this section describes how this task was undertaken.[8] We then assume that experts agree that these spot estimates of a_H, b_H and m_{kH} can be used for policy-making purposes. From this level of consensus regarding $f_H(x_H)$, we then calculate R_{lH} and R_{uH}.

4.1 ENPV

Under the ENPV model $x_H = \overline{r}_H$, the average risk-free rate over the horizon of the cash flow. There are a number of studies that have calibrated econometric models of Treasury bond yields in order to derive empirical schedules of the social discount rate using this ENPV approach; see, for example, Newell and Pizer (2003), Groom et al. (2007), Gollier et al. (2008), Hepburn et al. (2009), and Freeman et al. (2013). To estimate F_H here, we use the state-space model of Groom et al. (2007). We refer to this as model "GKPP". Our calibrations are taken directly from Groom et al. (2007).

Let $\theta_{ft} = \ln(100 r_{ft})$, with $r_{f0} = 4\%$. The variable θ_{ft} evolves according to:

$$\begin{aligned}\theta_{ft} &= \eta + \lambda_t \theta_{ft-1} + e_t \\ \lambda_t &= \eta_1 \lambda_{t-1} + u_t\end{aligned}$$

which is an AR(1) process with time-varying autoregressive parameter, λ_t. The error terms are independently and identically normally distributed (i.i.n.d.) with variance σ_e^2 and σ_u^2 respectively and zero means. To understand the properties of F_H in this case we run 25,000 simulations for $H \in [1, 400]$. In each simulation the parameters are drawn from the dis-

[8] In some cases F_H is supported on the interval $[-\infty, \infty]$ and therefore we cannot choose finite a_H, b_H to completely match this range. To overcome this problem, as described below, in such cases we choose a_H and b_H so that Prob$(x_H < a_H|F_H) = \varepsilon_a$ and Prob$(x_H > b_H|F_H) = \varepsilon_b$ for very small values of ε_a and ε_b. Our results are much more sensitive to the choice of a_H than b_H because $r_H \to r_{\min} = a_H$ as $H \to \infty$.

tributions estimated by Groom et al. (2007) before the AR(1) process is simulated for 400 periods. The support of F_H is parameterized by taking the observed minimum and maximum values for each H. $E_F[\exp(-Hx_H)]$ and m_{kH} are estimated by taking expectations across the 25,000 simulations for all values of H. More precise details of the simulation can be found in Appendix 2.

4.2 Consumption based asset pricing models

To derive a declining discount rate in a consumption-based asset pricing environment, we assume throughout that $\ln(c_t/c_{t-1}) = \mu_t + e_t$ where e_t is i.i.n.d. with mean zero and variance σ_t^2. We also assume that the process is homoskedastic with known variance, $\sigma_t = \sigma$, and that the parameters of the power utility function are $\gamma = 2$ and $\delta = 0$. The two models that we consider look at the parameterization of expected consumption growth, μ_t.

The first of this category are Markov regime switching models. Models of asset returns of this variety have been estimated by Cecchetti et al. (2000) for the US economy. Two states of the world are assumed to exist; good (G) and bad (B). If, at time $t-1$, the world is in state G (B) then $\mu_t = \mu_G$ (μ_B) with $\mu_G > \mu_B$. The Markov process is such that, conditional on being in state G (B) at time $t-1$, the probability of remaining in the same state at time t is $1 - \pi_G$ $(1 - \pi_B)$. Following Gollier (2012), we use two such models in order to simulate the uncertainty surrounding x_H. In the "Markov1" simulations, the good and bad state occur equally frequently and the latter reflects economic stagnation. In the "Markov2" simulations, the bad state occurs rarely but is a severe recessionary environment. The simulations are undertaken following the procedures described for the ENPV model, with the support, $E_F[\exp(-x_H H)]$ and the moments m_{kH} calculated analogously (see Appendix 2).

The third consumption based model incorporates parameter uncertainty; we call this "Param Uncert". This assumes that $\mu_t = \mu$ for all t, but that the value of μ is unknown. This approach is used as a justification for the French government's position on long-term social discounting, although the arguments are based on a simple numerical example rather than an empirical analysis; see (Lebegue 2005, p.102). To calibrate this model we follow Gollier (2012), and assume μ takes one of two values, $\mu_u > \mu_l$ with equal probability. In this case, the values of $E_F[\exp(-Hx_H)]$ and m_{kH} are known in closed form through the law of total probability; $E_F[h(x_H)] = 0.5E_F[h(x_H)|\mu = \mu_u] + 0.5E_F[h(x_H)|\mu = \mu_l]$ for any function $h(\cdot)$ with finite expectation. Further descriptions of the calibration are given in Appendix 2.

4.3 Heterogeneous agent models

There are two strands of heterogeneous agent model in the empirical literature on the term structure of discount rates. The first treats the individual agents, i, as countries or regions and uses this framework to derive a global social discount rate (Gollier et al. (2008); Gollier (2010); Emmerling (2010); Gollier (2012)). The second empirical approach, which we follow here, is to seek expert opinion on the values of x_H. Such a survey was the basis of the seminal 'gamma discounting' paper of Weitzman (2001) which reports $N = 2160$ survey responses, r_i, from professional economists to the question: '*Taking all relevant considerations into account, what real interest rate do you think should be used to discount over time the (expected) benefits and (expected) costs of projects being proposed to mitigate the possible effects of global climate change?*'. Implicit in this approach is the assumption that, for each expert, the discount rate is independent of the time horizon, $r_{iH} = r_i$ and therefore that the pdf $F_H(x_H)$ will also be $H-$independent.

Empirical schedules of the social discount rate following from this survey have been reported by Weitzman (2001), Jouini and Napp (2010) and Freeman and Groom (2013). Here we follow the original approach taken by Weitzman (2001). He argued that the present value of a future \$1 environmental benefit should be calculated using the sample frequency distribution of individual discount factors: $p_H = 2160^{-1}\sum_{i=1}^{2160}\exp(-Hr_i)$. While some have argued that using this method to justify a theory based on probability density functions conflates uncertainty and heterogeneity (e.g. Freeman and Groom (2013)), there remain some theoretical justifications for this approach. For example, it follows from equation 1 if q_i and c_{i0} are the same for all agents. If also follows from equation 2 if each agent is weighted equally: $z_i = 1/N$ for all i.

We estimate the sharp bounds for both the parametric and non-parametric descriptions of the data on expert opinions. The first we label as the "Gamma" model, because we follow Weitzman by using a gamma distribution to describe F_H for all H with $\alpha = 1.904$ and $\beta = 47.231$. This requires that we take $a_H = 0$, which is the lower support of the gamma distribution. We choose the upper bound, b_H by selecting the value that gives a 0.1% probability that a random variable drawn from $\Gamma(\alpha,\beta)$ will be greater than this value. This turns out to be $b_H = 19.12\%$. The benefit of the parametric assumption is that the moment generating function and the k^{th} non-central moments of a gamma distribution are known in closed form: $E_F[\exp(-Hx_H)] = (1 + H/\beta)^{-\alpha}$ and $m_{kH} = \beta^{-k}\prod_{i=0}^{k-1}(\alpha+i)$.

The second approach, which we call "Weit", is non-parametric and uses simply the sample frequency distribution of responses. This gives $a_H = -3\%$ and $b_H = 27\%$, which are the lowest and highest given answers to the survey question. The k^{th} non-central moment is now given by $m_{kH} = 2160^{-1} \sum_{i=1}^{2160} r_i^k$.

5 Results

The term structures of discount rates for the six calibrations that we consider, as given by $E_F [\exp(-Hx_H)]$, are illustrated by the solid lines in Figures 1 to 4 for $H \leq 400$ years. We call these the "baseline" calibrations in all cases.[9]

[Insert figures 1–4 around here]

As can be seen, these models all recommend a declining term structure of discount rates, which provides robust theoretical and empirical foundations for current policy choices in the UK, France and elsewhere. That said, there are also some clear discrepancies between the different figures. For example, for both GKPP and Markov2 the decline in the term structure occurs rapidly and then largely flattens out. For most of the other models the discount rate continues to noticeably decline throughout the horizon considered. The Weit model has negative values of R_H for large H, while all other models have positive discount rates for all horizons considered.

To understand what makes these term structures decline, Tables 1 reports the values of m_{kH} $(k \leq 4)$ for $H \in [1, 10, 50, 100, 400]$.

[Insert Table 1 around here]

These six models can be broadly divided into three groups. Group 1 contains GKPP and Markov1, Group 2 contains Param Uncert, Gamma and Weit, while Group 3's only element is Markov2.

The Group 1 models are distinctive; the first moment of x_H declines over time while the measures of spread (second to fourth non-central moments) are the lowest of all the models considered, particularly at longer horizons. In these cases the declining discount rates are primarily, but not exclusively, being driven by changes in the mean of the underlying economic process and not the Jensen's inequality effect that is usually used to justify DDRs.

Elements of Group 2, by contrast, have constant first moment across time apart from small fluctuations for Param Uncert caused by the use of a finite number of simulations.

[9] With the exception of the final graph, these term structures have been previously reported in the literature. The "GKPP" line can be compared against Figure 3 in Groom et al. (2007), the "Markov1" and "Markov2" lines can be respectively compared against the top line in Figure 5.3 of Gollier (2012) and the top graph in Figure 5.2 of the same source, the "Param Uncert" line can be compared against Figure 6.2 of Gollier (2012), while the "Gamma" line can be compared against Figure 1 in Freeman and Groom (2013).

The measures of spread for these calibrations are the greatest. Therefore, unlike Group 1, it is the uncertainty over x_H that drives the schedule of DDRs. Group 3 has properties that lie between those of Groups 1 and 2.

The support of the probability density functions, $[a_H, b_H]$ for $H \in [1, 10, 50, 100, 400]$, are reported in Table 2.

[Insert Table 2 around here]

The ENPV model is unusual in that the range $b_H - a_H$ gets wider as H increases. This is particularly caused by an increase in the upper bound over the first 50 years. For the consumption based models, the range $b_H - a_H$ gets narrower as H increases. This is primarily a consequence of time diversification in consumption risk. Because there is no certainty that future consumption must be above current consumption a_H is frequently negative in these cases, particularly for shorter horizons. For the heterogeneous agent models the assumption is that $f_H(x_H)$ is independent of H and therefore the values of a_H and b_H are fixed over time, as are the non-central moments reported in Table 1.

We would particularly draw the reader's attention to the difference between the Gamma and Weit calibrations as given by the solid lines in Figure 4. These term structures are noticeably different, particularly at long horizons, even though they are both based on the same survey data. From Table 1 we can see that the first four non-central moments are extremely similar for these two models and therefore they cannot explain the discrepancy between the two lines. Table 2, though, reports that $a_H = 0$ for Gamma while $a_H = -3\%$ for Weit at all horizons H. The gamma approximation effectively discards the zero and negative responses from the survey, while using the precise sample frequency does not. While only 46 (3) experts gave an answer of $r_i = 0\%$, $(r_i < 0\%)$ out of a total sample size of $N = 2160$, the treatment of these minority responses is crucial in determining the rate of decline of the social discount rate because $R_H \to \min\{r_i\} = a_H$ as $H \to \infty$.

In Table 3, we report an example of the discrete distributions that provide the sharp upper and lower bounds for R_H; v_{qlH} and v_{quH}. This is for the Markov1 model for horizons $H = 50$ years and $H = 200$ years. As described above, these bounds are determined by discrete distributions with non-zero mass in $[a_H, b_H]$ at between $(K+1)/2$ and $(K+3)/2$ points. The figures in square parentheses denote the probabilities associated with each non-zero mass point, π_{qlH} and π_{quH}.

[Insert Table 3 around here]

We calculate such discrete distributions for all models and all horizons $H \leq 400$ years and the sharp bounds on the discount rate are then derived from each. These are reported as the dotted lines in Figures 1–4.

Consider first the two models in Group 1. For GKPP, while the bounds are clearly distinct from the baseline calibration, the spread of potential values for R_H is not particularly great. This is because the declining schedule is driven by the mean, rather than the spread, and a_H remains well above zero for all horizons considered. For Markov1, while again the mean is the dominant effect reducing R_H with increasing H, the bounds spread is much wider than for GKPP. This is because a_H remains below zero for all H considered — see Table 2. For the three models in Group 2, the spread of possible values for R_H is wide for all H and K beyond approximately 75 years. For the Markov2 model, there is great uncertainty about the value of R_H at a horizon of approximately 100 years but there is more consensus for longer horizons. This is because a_H continues to increase rapidly even when H is large; see again Table 2.

Our central observation comes from these graphs. If there is a high level of agreement between economists about the partial characteristics of $f_H(x_H)$, which we would contend there is not, even then the rate with which the social discount rate declines can take a wide range of potential values. This will have obvious implications for the extent to which conclusions can be drawn from CBA of intergenerational projects.

To illustrate this point, in Table 4 we calculate a Social Cost of Carbon (SCC) in US dollars per ton of carbon ($/tC) from each of the term structures reported in Figures 1–4 using the schedule of marginal carbon damages provided by Newell and Pizer (2003). These cash flows have a horizon of 400 years, with 50% of the undiscounted costs arising by year 170.

[Insert Table 4 around here]

Excluding Weit, the range of estimates for the baseline calibrations range from 8.3$/tC to 32.0$/tC. Weit is a considerable outlier, with an estimated SCC of $979.8/tC. This results from the fact that R_H is negative under this model for far horizons. As a consequence very long term environmental damages are compounded rather than discounted. Again, this shows the sensitivity of the results to the choice of a_H for high values of H, which is the primary feature that distinguishes the Weit model from the Gamma model.

For some models, provided $K = 4$, it is possible to estimate the SCC relatively accurately. For example, the GKPP (Param Uncert) model has a range of $16.3/tC ($13.6/tC) to 19.9/tC ($16.5/tC). For Markov2, the range is $8.2/tC to $9.2/tC. For many models, though, great uncertainty remains over the Social Cost of Carbon even when we assume that moments of $f_H(x_H)$ are perfectly known up to and including the kurtosis. Under Gamma (Markov1), the range is $13.6/tC ($25.2/tC) to $46.1/tC ($85.0/tC). For values of $K < 4$, the ranges increase substantially. For example, if we only know the mean, variance and skewness of $f_H(x_H)$ under the Markov1 model, then the SCC can lie anywhere in the range \$22.6/tC – \$242.2/tC.

Again, the Weit model is the outlier, with a range of potential values for the SCC of \$11.6/tC – \$84,592/tC when $K = 4$. This schedule would make it almost impossible for a policy maker to make any serious economic decisions about a wide range of climate change mitigation programmes. For $K = 1$, we cannot even preclude the value of one and a half million dollars per ton of carbon under this calibration.

Tables 5 reports the present value of the official estimates of benefits from Phase 1 of the HS2 rail link (London to Birmingham) that is currently being considered in the UK. The cash flows are taken directly from the HS2 official website. These arise over a 75 year period to 2085, with 50% of the undiscounted benefits occurring by year 53. Table 6 considers instead the estimated costs of decommissioning nineteen nuclear power stations in the UK as given in the Nuclear Decommissioning Authority report and account 2012/13. While these span over a longer time horizon than HS2 (125 years), the half-life, as it were, is shorter, with 50% of the undiscounted costs occurring by year 29. Further details on these cash flow estimates are available on request from the authors.

[Insert Table 5 & 6 around here]

As might be expected, given the shorter time horizons involved, the present value bounds are narrower than for the social costs of carbon. Nevertheless, particularly when $K \leq 3$, the uncertainty is still of policy importance. For example, using gamma discounting and $K = 3$, the NPV of the benefits of HS2 lies between £20.7bn–£28.3bn and the present value costs of decommissioning the previous generation of nuclear power stations lies in the interval £44.8bn–£51.3bn. With lower K, this uncertainty is widened considerably further.

6 Conclusion

There are strong theoretical arguments for using a declining term structure of discount rates for intergenerational projects, and these are largely based on uncertainty about future growth and interest rates. So persuasive have these arguments been that they are now recognised in government policies and recommendations in the UK, France, Norway and the US. The practical question that necessarily follows is: how can uncertainty be characterised and the theory be operationalised?

So far policy makers have sought a range of expert economists' advice on the best route forward, and the experts have provided them with empirical estimates of the certainty equivalent term structure of social discount rates based on specific theories and characterisations of uncertainty; see, for example, Gollier (2012), Arrow et al. (2012).

The fact of the matter is, though, that we know far less about the future state of the world and the uncertainty surrounding growth and the interest rate than the complete characterisations of uncertainty that many of these models suggest. This raisies the question, how certain are we about the certainty equivalent social discount rate?

Starting from a position of relative ignorance about our knowledge of uncertainty in the future, this paper shows that, even if these experts have strongly overlapping views on the primatives of social discounting, and agree on the first four moments of the distribution of growth or the interest rate, the empirical term structure of the certainty equivalent social discount rate emerging from many theoretical models cannot be positioned within anything but very wide bounds, particularly for long time horizons. The obvious implication of this is that policy prescriptions become less crisp for all but the highest return public projects.

For instance viable estimates for the Social Cost of Carbon and the Net Present Values of other intergenerational project become alarming dispersed from this position of ignorance. Apparently trivial disagreements over parameterization choices can lead to significant differences in policy recommendations.[10]

While not quite a dismal theorem, overall this paper presents a depressing empirical finding for practitioners of Cost Benefit Analysis. Even if we are willing to put issues of Knightian uncertainty to one side, we must accept that we know little about the true nature of uncertainty in the distant future. This admission of ignorance means that estimated present values are likely to be so imprecise as to provide only minimal guidance to policy makers on intergenerational projects. We may have to look elsewhere for a decision making apparatus in such cases.

References

Almond, D. (2006), 'Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 U.S. population', *Journal of Political Economy* **114**, 672–

[10] As a further example of this point that does not invoke sharp bounds on R_H, we note that Weitzman (2001) chose a gamma distribution to characterize his survey data largely because it has a well-known moment generating function, leading to an elegant closed form expression for the $H-$period social discount rate. He might, instead, have chosen a Wald (Inverse Gaussian) distribution, which is also supported on $[0,\infty)$ and leads to a closed form expression for R_H. In this Wald distribution case, the estimated SCC is \$14.6/tC, more than 25% lower than the baseline Gamma estimate of \$20.0/tC. Indeed, a Kolmogorov-Smirnov test marginally prefers the Wald distribution to the Gamma distribution when fitting the survey data. Details of the properties of R_H when $F_H(x_H)$ is given by a Wald distribution are available on request from the authors.

712.

Ang, A. & Liu, J. (2004), 'How to discount cashflows with time-varying expected returns', *Journal of Finance* **59**, 2745–2783.

Arrow, K., Cropper, M., Gollier, C., Groom, B., Heal, G., Newell, R., Nordhaus, W., Pindyck, R., Pizer, W., Portney, P., Sterner, T., Tol, R. & Weitzman, M. (2012), How should benefits and costs be discounted in an intergenerational context? The views of an expert panel. Resources for the Future Discussion Paper, RFF DP 12-53.

Arrow, K., Cropper, M., Gollier, C., Groom, B., Heal, G., Newell, R., Nordhaus, W., Pindyck, R., Pizer, W., Portney, P., Sterner, T., Tol, R. & Weitzman, M. (2013), 'Determining benefits and costs for future generations', *Science* **341**, 349–350.

Cecchetti, S. G., Lam, P. & Mark, N. C. (2000), 'Asset pricing with distorted beliefs: Are equity returns too good to be true?', *American Economic Review* **90**, 787–805.

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

Eckberg, A. E. (1977), 'Sharp bounds on Laplace-Stieltjes transforms, with applications to various queueing problems', *Mathematics of Operations Research* **2**, 135–142.

Emmerling, J. (2010), Discounting and intragenerational equity. Mimeo: Toulouse School of Economics.

Freeman, M. C. (2010), 'Yes, we should discount the far-distant future at its lowest possible rate: A resolution of the Weitzman Gollier puzzle', *Economics: The Open-Access, Open-Assessment E-Journal* **4**(2010–13).

Freeman, M. C. & Groom, B. (2013), Positively gamma discounting. Loughborough University Working Paper.

Freeman, M. C., Groom, B., Panopoulou, E. & Pantelidis, T. (2013), Declining discount rates and the Fisher Effect: Inflated past, discounted future? Grantham Research Institute on Climate Change and the Environment Working Paper No. 109, London School of Economics.

Gollier, C. (2009), 'Should we discount the far-distant future at its lowest possible rate?', *Economics - The Open-Access, Open-Assessment E-Journal* **3**(2009–25), 1–14.

Gollier, C. (2010), Discounting, inequalities and economic convergence. CESifo working paper: Industrial Organisation, No. 3262.

Gollier, C. (2012), *Pricing the Planet's Future: The Economics of Discounting in an Uncertain World*, Princeton University Press, Princeton.

Gollier, C., Koundouri, P. & Pantelidis, T. (2008), 'Declining discount rates: Economic justifications and implications for long-run policy', *Economic Policy* **23**(56), 757–795.

Gollier, C. & Weitzman, M. L. (2010), 'How should the distant future be discounted when discount rates are uncertain?', *Economics Letters* **107**(3), 350–353.

Gollier, C. & Zeckhauser, R. (2005), 'Aggregation of heterogeneous time preferences', *Journal of Political Economy* **113**, 878–896.

Groom, B., Koundouri, P., Panopoulou, E. & Pantelidis, T. (2007), 'Discounting the distant future: How much does model selection affect the certainty equivalent rate?', *Journal of Applied Econometrics* **22**(3), 641–656.

Heal, G. (2012), Discounting - uncertainty, disagreement or both? Mimeo; Columbia Business School.

Heal, G. & Millner, A. (2013), Discounting under disagreement. Working paper, Grantham Research Institute, No. 112.

Hepburn, C., Koundouri, P., Panopoulou, E. & Pantelidis, T. (2009), 'Social discounting under uncertainty: A cross-country comparison', *Journal of Environmental Economics and Management* **57**(2), 140–150.

Iverson, T. (Forthcoming), 'Minimax regret discounting', *Journal of Environmental Economics and Management* .

Johnson, M. A. & Taaffe, M. R. (1993), 'Tchebycheff systems for probabilistic analysis', *American Journal of Mathematical and Management Sciences* **13**, 83–111.

Jouini, E., Marin, J.-M. & Napp, C. (2010), 'Discounting and divergence of opinion', *Journal of Economic Theory* **145**, 830–859.

Jouini, E. & Napp, C. (2010), Aggregation of discount rates: An equilibrium approach. Mimeo; Université Paris-Dauphine.

Karlin, S. & Studden, W. (1966), *Tchebycheff Systems: With Applications in Analysis and Statistics*, John Wiley & Sons, New York.

Knight, F. H. (1921), *Risk, Uncertainty and Profit*, Hart, Schaffner & Marx; Houghton Mifflin Company, Boston MA.

Lebegue, D. (2005), Revision du taux de actualisation des investissem- net publics. rapport du groupe de experts, commisariat generale de plan., Technical report, Commisariat Generale de Plan.

Li, C.-Z. & Löfgren, K.-G. (2000), 'Renewable resources and economic sustainability: A dynamic analysis with heterogeneous time preferences', *Journal of Environmental Economics and Management* **40**, 236–250.

Newell, R. G. & Pizer, W. A. (2003), 'Discounting the distant future: How much do uncertain rates increase valuations?', *Journal of Environmental Economics and Management* **46**, 52–71.

Traeger, C. P. (Forthcoming), 'What's the rate: Disentangling the Weitzman and Gollier effect', *Journal of Environmental Economics and Management* .

Weitzman, M. L. (1998), 'Why the far-distant future should be discounted at its lowest possible rate', *Journal of Environmental Economics and Management* **36**(3), 201–208.

Weitzman, M. L. (2001), 'Gamma discounting', *American Economic Review* **91**, 260–271.

Weitzman, M. L. (2009), 'On modeling and interpreting the economics of catastrophic climate change', *The Review of Economics and Statistics* **91**(1), 1–19.

7 Appendix 1

In this appendix, we describe in more detail the method that we use for determining R_{uH} and R_{lH}. When $f_H(x_H)$ is supported on the interval $[a_H, b_H]$ for $a_H \neq 0$, we first undertake a change of variables. Define $y_H = x_H - a_H$, $g^*_H(y_H) = g_H(x_H)$ and $\Im^*_{KH}$ as the set containing all elements $g^*_H(y_H)$. The lower bounds for R_H are derived from the bounds for $E_{g^*}[\exp(-Hy_H)]$.

$$R_{lH} = -\frac{1}{H}\ln\left(\exp(-Ha_H)\sup\left[E_{g^*}[\exp(-Hy_H)]\,|g^*_H \in \Im^*_{KH}\right]\right)$$

and there is an analogous expression for R_{uH}, with the supremum replaced by the infimum.

Well defined probability density functions are elements of $\Im^*_{KH}$ if and only if they are supported on the interval $[0, B_H]$, where $B_H = b_H - a_H$ and if they have the same K non-central moments, M_{kH} for $k \leq K$. These non-central moments are given by the binomial theorem:

$$M_{kH} = \sum_{\zeta=0}^{k} (-1)^{\zeta} \begin{pmatrix} k \\ \zeta \end{pmatrix} m_{kH}^{k-\zeta} a_H^{\zeta}$$

We then use the method described in the body of the paper to find $\sup\left[E_{g^*}[\exp(-Hy_H)]\,|g^*_H \in \Im^*_{KH}\right]$ and $\inf\left[E_{g^*}[\exp(-Hy_H)]\,|g^*_H \in \Im^*_{KH}\right]$ and then note that once the values of V_{qlH}, V_{quH}, π_{qlH} and π_{quH} have been determined, R_{lH} is then given by:

$$\begin{aligned} R_{lH} &= -\frac{1}{H}\ln\left(\exp(-Ha_H)\sup\left[E_{g^*}[\exp(-Hy_H)]\,|g^*_H \in \Im^*_{KH}\right]\right) \\ &= -\frac{1}{H}\ln\left(\sum_{q=1}^{\varpi} \pi_{quH}\exp(-Hv_{quH})\right) \end{aligned}$$

where $v_{qlH} = V_{qlH} + a_H$. An analogous expression follows for R_{uH}.

Closed form solutions for V_{qlH}, V_{quH}, π_{qlH} and π_{quH} are available for $K \leq 3$. When $K = 1$, the lower bound has only one mass point which is on the mean value; $V_{1lH} = M_{1H}$. The upper bound has mass at $V_{1uH} = 0$ and $V_{2uH} = B_H$ only and the probability π_{1uH} is set to ensure that the mean is equal to M_{1H}; $\pi_{1uH} = (B_H - M_{1H})\,/\,B_H$. Closed form solutions for $K = 2$ and $K = 3$ are given in Eckberg (1977). For the case $K = 2$:

	$q=1$	$q=2$
V_{qlH}	$\frac{M_{1H}B_H - M_{2H}}{B_H - M_{1H}}$	B_H
π_{qlH}	π_1	$1-\pi_1$

	$q=1$	$q=2$
V_{quH}	0	$\frac{M_{2H}}{M_{1H}}$
π_{quH}	$\frac{M_{2H} - M_{1H}^2}{M_{2H}}$	$\frac{M_{1H}^2}{M_{2H}}$

where $\pi_1 = (B_H - M_{1H})^2 / \left(M_{2H} - M_{1H}^2 + (B_H - M_{1H})^2\right)$. For $K = 3$:

	$q = 1$	$q = 2$
V_{qlH}	$\frac{A_1 - \chi}{2}$	$\frac{A_1 + \chi}{2}$
π_{qlH}	$\frac{\chi + A_1 - 2M_{1H}}{2\chi}$	$\frac{\chi - A_1 + 2M_{1H}}{2\chi}$

	$q = 1$	$q = 2$	$q = 3$
V_{quH}	0	ζ	B_H
π_{quH}	$1 - \pi_1^* - \pi_2^*$	π_1^*	π_2^*

where:

$$A_1 = \frac{M_{3H} - M_{1H}M_{2H}}{M_{2H} - M_{1H}^2} \qquad A_2 = \frac{M_{2H}^2 - M_{1H}M_{3H}}{M_{2H} - M_{1H}^2}, \qquad \chi = \sqrt{A_1^2 + 4A_2}$$

$$\pi_2^* = \frac{M_{1H}M_{3H} - M_{2H}^2}{M_{1H}B_H^3 - 2M_{2H}B_H^2 + M_{3H}B_H}, \qquad \pi_1^* = \frac{(M_{1H} - q_2 B_H)^2}{M_{2H} - q_2 B_H^2}, \qquad \zeta = \frac{M_{2H} - q_2 B_H^2}{M_{1H} - q_2 B_H}$$

8 Appendix 2

Here we describe in detail the way in which the simulations are undertaken.

8.1 'GKPP' State Space model:

The parameter estimates (with associated standard errors) are $\eta = 0.510$ (0.0082), $\eta_1 = 0.990$ (0.002), $\ln(\sigma_e^2) = -9.158$ (1.324), $\ln(\sigma_u^2) = -6.730$ (0.144). To characterize F_H we run 25,000 simulations for $H \in [1, 400]$. For each simulation we initially draw values of η, η_1, $\ln(\sigma_e^2)$ and $\ln(\sigma_u^2)$ at random from normal distributions with the appropriate mean and variance. Following Groom et al. (2007) we discard simulations with $\eta_1 > 1$. We then simulate the AR(1) process for 400 periods. This process can lead to some high values of r_{ft}, particularly when η_1 is close to 1, and therefore we discard simulations where $r_{ft} > 100\%$ at any point in the 400 year horizon. In order to parameterize a_H and b_H, we note the minimum and maximum values taken by x_H across the 25,000 simulations for all H. We then approximate a functional form: $a_H = \phi_a(\min(x_H), H)$, $b_H = \phi_b(\max(x_H), H)$ to smooth small fluctuations caused by the fact that $\min(x_H)$ and $\max(x_H)$ are estimated over a finite number of simulations. The values of $E_F[\exp(-Hx_H)]$ and m_{kH} are calculated by taking expectations across the 25,000 simulations for all values of H.

8.2 Consumption Based models

8.2.1 Markov 1 and Markov 2:

For each simulation, we set the initial state to G. For model “Markov1”, $\mu_G = 2\%$, $\mu_B = 0\%$, $\pi_G = \pi_B = 1\%$ and $\sigma_t = \sigma = 3.6\%$ while for "Markov 2”, $\mu_G = 2.25\%$, $\mu_B = -6.78\%$, $\pi_G = 2.2\%$, $\pi_B = 48.4\%$ and $\sigma_t = \sigma = 3.13\%$.

Following the techniques used for model GKPP, we estimate the properties of F_H across 25,000 simulations. This enables us to calculate both $E_F\left[\exp\left(-Hx_H\right)\right]$ and the moments of x_H. To parameterize a_H and b_H we again note the minimum and maximum values taken by x_H across the 25,000 simulations for all H and use smoothing functions, $a_H = \phi_a^*\left(\min\left(x_H\right), H\right)$, $b_H = \phi_b^*\left(\max\left(x_H\right), H\right)$, to eliminate small fluctuations caused by the use of a finite number of simulations.

8.2.2 "Param Uncert":

The calibration is based upon Gollier (2012); $\mu_u = 3\%$ and $\mu_l = 1\%$. These outcomes are equally likely. The limits of the support, a_H and b_H, are estimated by setting $\Phi_l\left(a_H\right) = 0.1\%$ and $\Phi_u\left(b_H\right) = 99.9\%$, where Φ_u and Φ_l denote the cumulative density functions of normal variables $N\left(\mu_u, \sigma^2\right)$ and $N\left(\mu_l, \sigma^2\right)$ respectively.

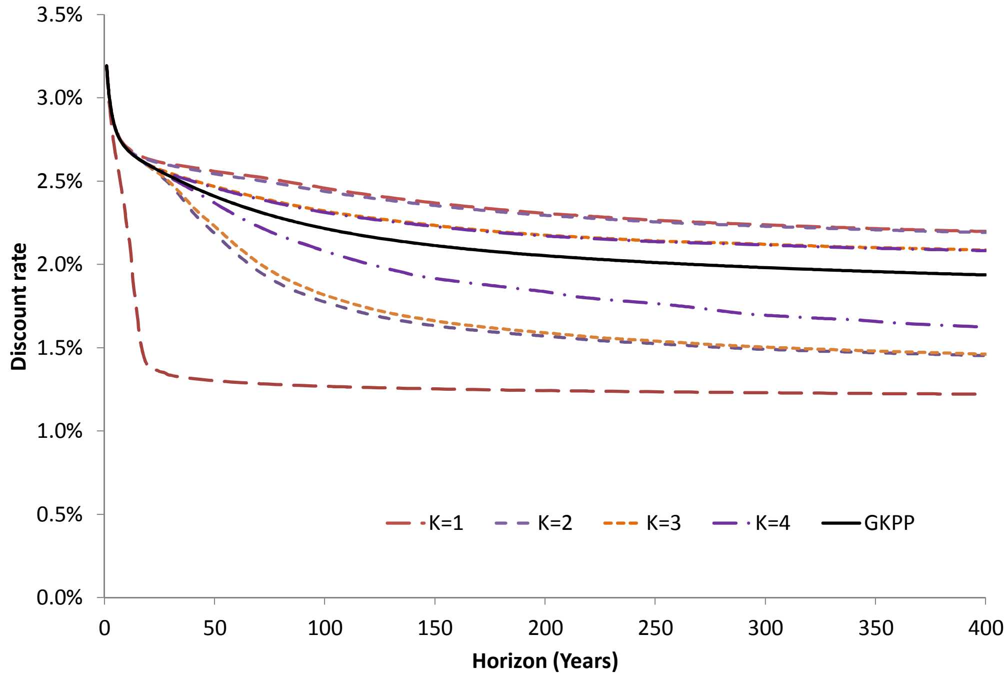
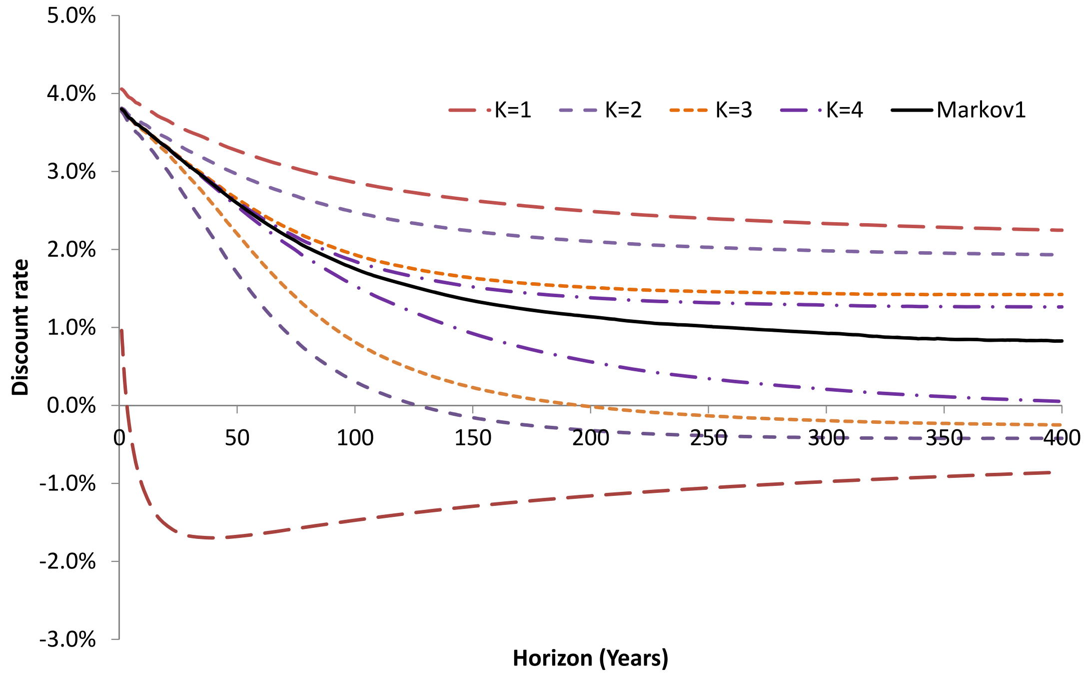


Figure 1. This figure presents the term structure of social discount rates as given through the ENPV setting by the state-space model of Groom et al. (2007). The solid line is the baseline parameterization from this model, which can be compared against Figure 3 in Groom et al. (2007). We then present upper and lower bounds on R_H conditional on matching the first K moments of $F_H(x_H)$ for $K \in [1, 4]$ as reported in Table 1. The support of the probability density function is also restricted to lie in $[a_H, b_H]$ for values of a_H and b_H reported in Table 2.



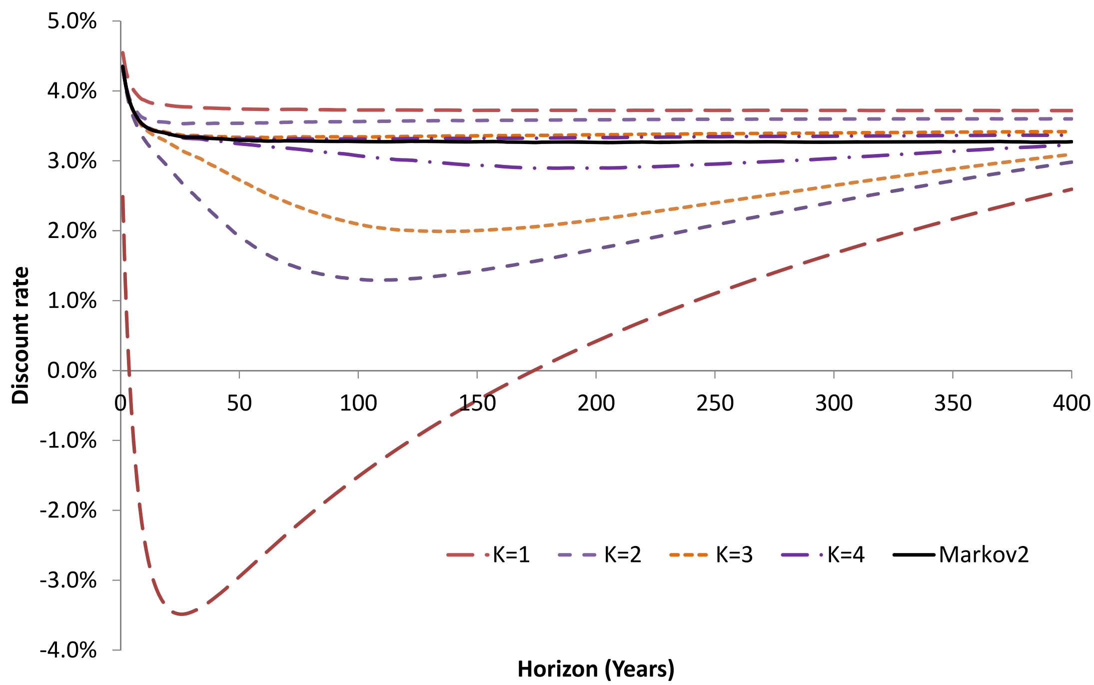


Figure 2. As Figure 1, except $F_H(x_H)$ is now derived from two Markov switching models for logarithmic consumption growth. The solid line in the top graph can be compared against the top line in Figure 5.3 of Gollier (2012), while the solid line in the bottom graph should be compared against the top graph in Figure 5.2 of the same source.

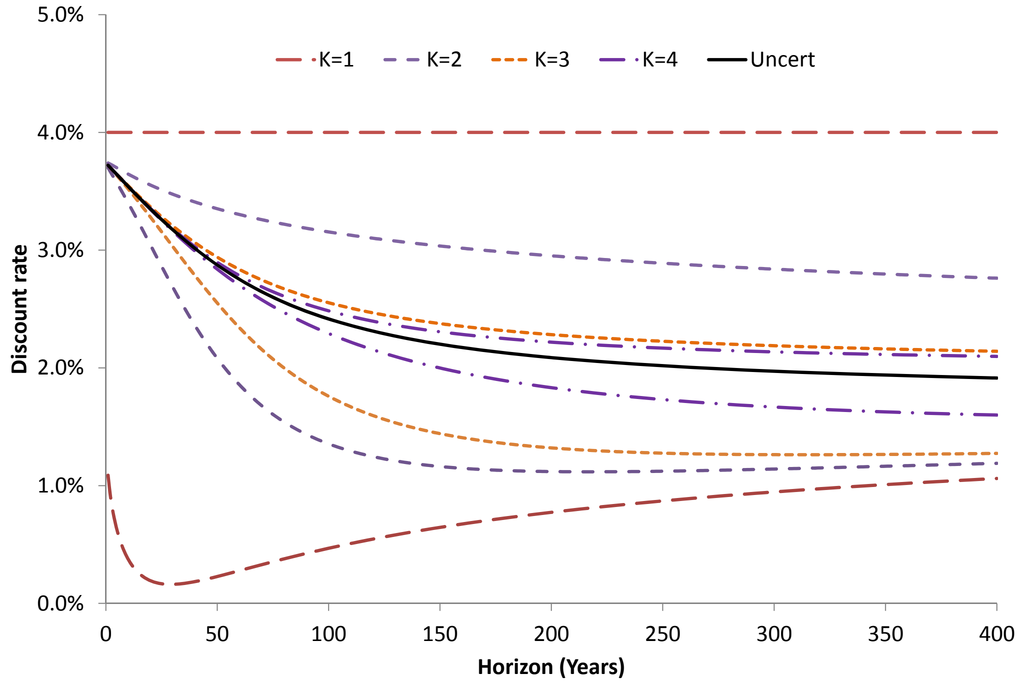
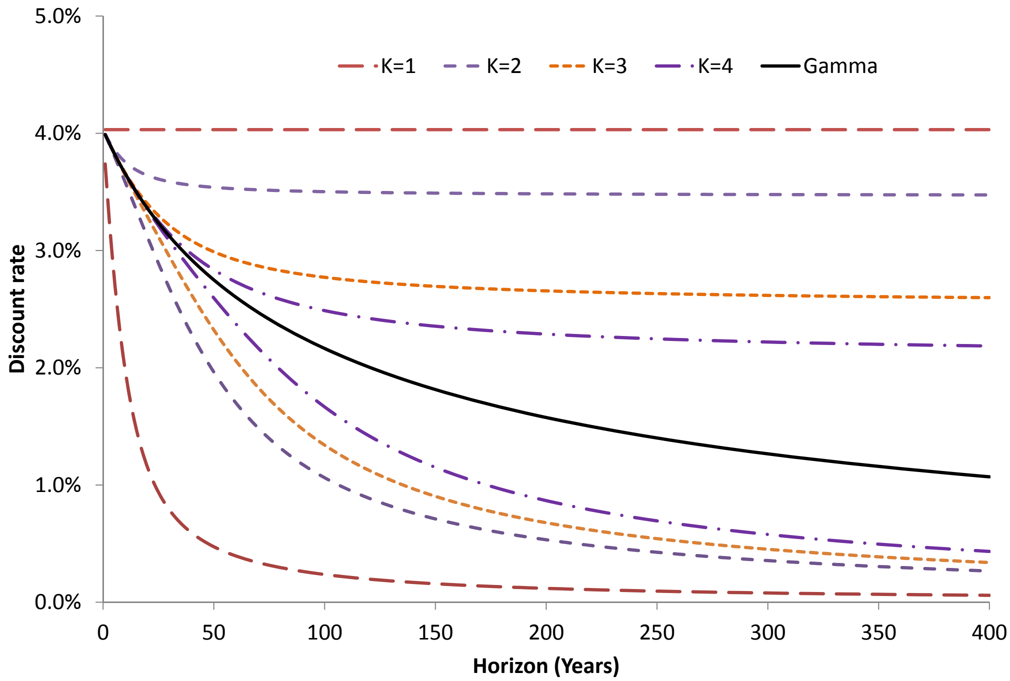


Figure 3. As Figures 1 and 2, except $F_H(x_H)$ is now derived from the parameter uncertainty model for logarithmic consumption growth. The solid line can be compared against Figure 6.2 of Gollier (2012).



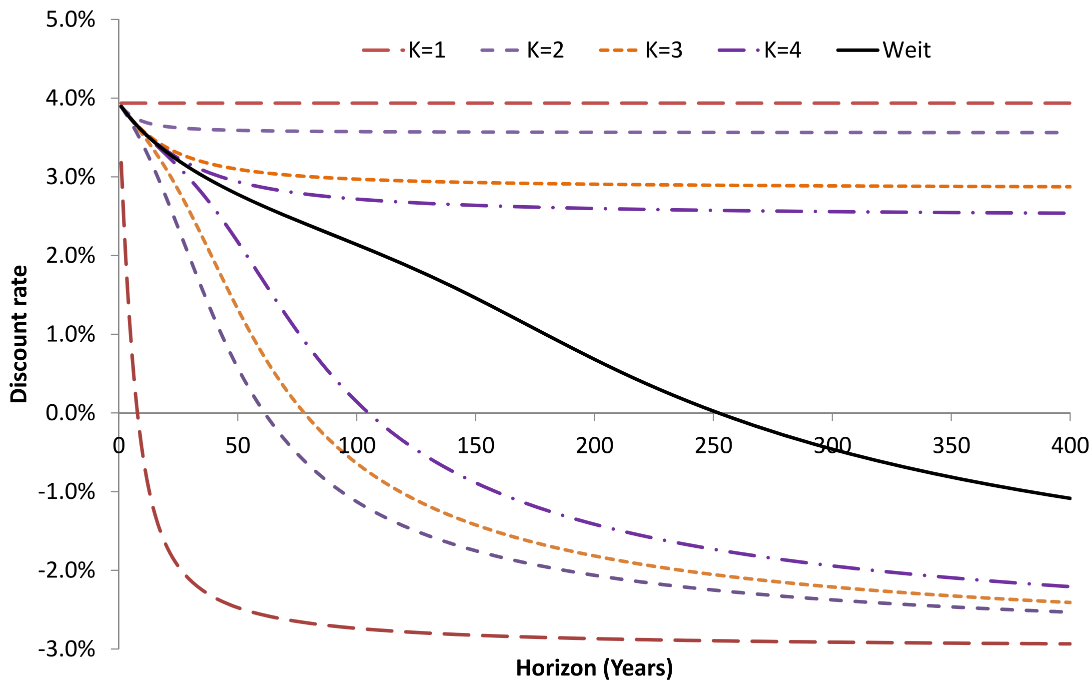


Figure 4. As Figures 1–3, except $F_H(x_H)$ is now derived from Weitzman's seminal (2001) "gamma discounting" survey. The solid line in the top graph can be compared against Figure 1 of Freeman and Groom (2013). The top graph follows Weitzman in using a gamma distribution to approximate the sample frequency of responses, while the bottom graph does not make this approximation and uses the raw sample distribution instead.

	GKPP	Markov1	Markov2	Param Uncert	Gamma	Weit
			Panel A: m_{1H}; First non-central moment			
H=1	3.192%	4.057%	4.546%	4.030%	4.031%	3.937%
H=10	2.707%	3.830%	3.869%	4.016%	4.031%	3.937%
H=50	2.560%	3.267%	3.741%	4.021%	4.031%	3.937%
H=100	2.458%	2.859%	3.727%	4.014%	4.031%	3.937%
H=400	2.198%	2.247%	3.718%	4.015%	4.031%	3.937%
			Panel B: m_{2H}; Second non-central moment			
H=1	0.103%	0.679%	0.603%	0.735%	0.248%	0.242%
H=10	0.076%	0.202%	0.217%	0.254%	0.248%	0.242%
H=50	0.081%	0.130%	0.155%	0.212%	0.248%	0.242%
H=100	0.079%	0.102%	0.147%	0.206%	0.248%	0.242%
H=400	0.055%	0.060%	0.140%	0.203%	0.248%	0.242%
			Panel C: m_{3H}; Third non-central moment			
H=1	3.344E-05	6.918E-04	6.415E-04	7.589E-04	2.048E-04	2.241E-04
H=10	2.194E-05	1.192E-04	1.201E-04	1.769E-04	2.048E-04	2.241E-04
H=50	5.186E-05	5.589E-05	6.858E-05	1.260E-04	2.048E-04	2.241E-04
H=100	5.281E-05	3.951E-05	6.021E-05	1.191E-04	2.048E-04	2.241E-04
H=400	1.788E-05	1.763E-05	5.362E-05	1.145E-04	2.048E-04	2.241E-04
			Panel D: m_{4H}; Fourth non-central moment			
H=1	1.098E-06	1.336E-04	1.012E-04	1.574E-04	2.127E-05	2.923E-05
H=10	6.616E-07	7.845E-06	8.235E-06	1.386E-05	2.127E-05	2.923E-05
H=50	8.736E-06	2.551E-06	3.183E-06	7.958E-06	2.127E-05	2.923E-05
H=100	7.786E-06	1.621E-06	2.559E-06	7.250E-06	2.127E-05	2.923E-05
H=400	9.660E-07	5.564E-07	2.076E-06	6.776E-06	2.127E-05	2.923E-05

Table 1. This table presents the first four moments of $F_H(x_H)$ as estimated from the different models for $H \in [1, 10, 50, 100, 400]$.

	GKPP	Markov1	Markov2	Param Uncert	Gamma	Weit
			Panel D: Lower support of pdf, a_H			
H=1	2.17%	-29.49%	-19.14%	-20.25%	0.00%	-3.00%
H=10	1.32%	-8.20%	-10.85%	-5.04%	0.00%	-3.00%
H=50	1.28%	-3.35%	-5.06%	-1.15%	0.00%	-3.00%
H=100	1.26%	-2.28%	-2.56%	-0.22%	0.00%	-3.00%
H=400	1.22%	-1.05%	2.43%	0.89%	0.00%	-3.00%
			Panel E: Upper support of pdf, b_H			
H=1	4.77%	21.74%	21.67%	28.25%	19.12%	27.00%
H=10	11.15%	12.34%	12.44%	13.04%	19.12%	27.00%
H=50	100.00%	8.30%	8.44%	9.15%	19.12%	27.00%
H=100	100.00%	7.00%	7.14%	8.22%	19.12%	27.00%
H=400	100.00%	4.98%	5.11%	7.11%	19.12%	27.00%

Table 2. This table presents the upper and lower support levels, a_H and b_H, for the probability density functions that we consider when generating sharp upper and lower bounds for R_H for $H \in [1, 10, 50, 100, 400]$.

	Lower bound on R_H			Upper bound on R_H		
	v_{1uH}	v_{2uH}	v_{3uH}	v_{1lH}	v_{2lH}	v_{3lH}
			Panel A: $H = 50$			
K=1	-3.35%	8.30%		3.27%		
	[43.22%]	[56.78%]		[100%]		
K=2	-3.35%	3.62%		2.80%	8.30%	
	[5.11%]	[94.89%]		[91.49%]	[8.51%]	
K=3	-3.35%	3.21%	8.30%	1.22%	4.42%	
	[2.74%]	[92.70%]	[4.56%]	[36.13%]	[63.87%]	
K=4	-3.35%	1.78%	4.66%	0.79%	4.07%	8.30%
	[0.69%]	[46.54%]	[52.77%]	[25.99%]	[72.85%]	[1.16%]
			Panel B: $H = 200$			
K=1	-1.55%	5.90%		2.49%		
	[45.82%]	[54.18%]		[100%]		
K=2	-1.55%	2.87%		2.04%	5.90%	
	[8.56%]	[91.44%]		[88.43%]	[11.57%]	
K=3	-1.55%	2.44%	5.90%	1.11%	3.60%	
	[4.62%]	[88.77%]	[6.61%]	[44.67%]	[55.33%]	
K=4	-1.55%	1.49%	3.80%	0.85%	3.25%	5.90%
	[1.35%]	[53.70%]	[44.96%]	[34.07%]	[63.90%]	[2.03%]

Table 3. For model Markov1 and $H = 50, 200$ years, this table presents the discrete distributions that result in the upper and lower bounds for R_H; v_{quH} and v_{qlH}. Figures in square parentheses give the probabilities, π_{quH} and π_{qlH}, associated with each value of x_H that has non-zero mass.

	GKPP	Markov1	Markov2	Param Uncert	Gamma	Weit
			Panel A: Baseline SCC			
	17.5	32.0	8.3	14.4	20.0	979.8
			Panel B: Lowest SCC			
K=1	14.7	10.3	6.3	5.4	5.3	5.6
K=2	14.9	13.8	7.1	8.6	7.2	6.9
K=3	16.2	22.6	8.0	12.9	11.1	9.9
K=4	16.3	25.2	8.2	13.6	13.6	11.6
			Panel C: Highest SCC			
K=1	42.3	2,452.9	471.6	89.2	188.9	1,549,070
K=2	25.6	447.0	32.4	38.9	83.7	309,301
K=3	24.7	242.2	17.1	27.3	63.8	189,093
K=4	19.9	85.0	9.2	16.5	46.1	84,592

Table 4. This table presents estimates of the Social Cost of Carbon (SCC) in terms of dollars per tonne of carbon ($/tC) for each model considered. Marginal damages of carbon emissions are taken from Newell and Pizer (2003). Panel A presents the estimate of the SCC derived from each schedule of discount rates. Panels B and C present minimum and maximum estimates of the SCC for each model, where the discount rate is respectively taken from the sharp upper and lower bounds of R_H for that model.

	GKPP	Markov1	Markov2	Param Uncert	Gamma	Weit
			Panel A: Baseline HS2 benefits			
	27.0	25.0	18.4	21.9	23.1	22.9
			Panel B: Lowest HS2 benefits			
K=1	25.3	18.5	15.4	13.9	13.8	14.3
K=2	25.5	21.1	16.8	18.0	16.7	16.4
K=3	26.3	24.3	18.1	21.3	20.7	19.9
K=4	26.4	24.7	18.3	21.7	22.1	21.2
			Panel C: Highest HS2 benefits			
K=1	44.7	199.8	351.8	74.1	66.7	330.4
K=2	29.9	39.6	33.7	31.3	33.2	73.5
K=3	29.4	30.9	23.5	25.4	28.3	50.8
K=4	27.6	25.7	18.9	22.3	25.0	32.8

Table 5. As Table 4, but now this reflects the present value of the estimated benefits of Phase 1 of the High Speed 2 (HS2) rail line in the UK; London to Birmingham (£bn).

	GKPP	Markov1	Markov2	Param Uncert	Gamma	Weit
			Panel A: Baseline decommissioning costs			
	51.9	48.7	43.4	46.1	46.8	46.9
			Panel B: Lowest decommissioning costs			
K=1	50.6	42.9	40.4	39.1	38.9	39.5
K=2	50.7	45.3	42.0	42.9	41.7	41.6
K=3	51.4	48.0	43.2	45.5	44.8	44.4
K=4	51.4	48.3	43.3	45.9	45.9	45.4
			Panel C: Highest decommissioning costs			
K=1	67.0	209.4	314.2	94.1	85.3	376.7
K=2	54.1	62.7	55.5	53.8	55.2	102.6
K=3	53.7	54.6	47.5	49.0	51.3	78.5
K=4	52.3	49.4	43.8	46.4	48.6	59.0

Table 6. As Tables 4 & 5, but now the present values are associated with the costs of decommissioning the previous generation of nuclear power stations in the UK as given in the 2012/13 report and accounts of the Nuclear Decommissioning Authority (£bn).