Discounting the distant future: How much does model selection affect the certainty equivalent rate?

Ben Groom[*] Phoebe Koundouri[†] Ekaterini Panopoulou[‡]
Theologos Pantelidis[§]

December 20, 2004

Abstract

Evaluating investments with long-term consequences using discount rates that decline with the time horizon, (Declining Discount Rates or DDRs) means that future welfare changes are of greater consequence in present value terms. Recent work in this area has turned towards operationalising the theory and establishing a schedule of DDRs for use in cost benefit analysis. Using US data we make the following points concerning this transition: i) model selection has important implications for operationalising a theory of DDRs that depends upon uncertainty; ii) misspecification testing naturally leads to employing models that account for changes in the interest rate generating mechanism. Lastly, we provide an analysis of the policy implications of DDRs in the context of climate change for the US and show that the use of a state space model can increase valuations by 150% compared to conventional constant discounting.

JEL classification: *C13, C53, Q2, Q4*

Keywords: *long-run discounting, interest rate forecasting, state-space models, regime-switching models, climate change policy.*

---

[*]Department of Economics, University College London.

[†]Department of Economics, University of Reading, UK and Department of Economics, University College London, UK.

[‡]Department of Banking and Financial Management, University of Piraeus, Greece and Department of Economics, National University of Ireland Maynooth. *Correspondence to:* Ekaterini Panopoulou, Department of Economics, National University of Ireland Maynooth, Co.Kildare, Republic of Ireland. E-mail: apano@may.ie. Tel: 00353 1 7083793. Fax: 00353 1 7083934.

[§]Department of Banking and Financial Management, University of Piraeus, Greece.
Acknowledgements: We are grateful to Christian Gollier, Cameron Hepburn, Dimitrios Malliaropulos, David Pearce, Nikitas Pittis, participants in the 2003 Royal Economic Society Conference, the 2004 Applied Environmental Economics Conference Royal Society, the First Hispanic Portuguese Congress of Environmental and Natural Resource Economics, the 13th Annual EAERE Conference, and seminar participants at the University College London and Reading University for helpful comments and suggestions. Panopoulou and Pantelidis thank the EU for financial support under the "PYTHAGORAS: Funding of research groups in the University of Piraeus" through the Greek Ministry of National Education and Religious Affairs.

1 Introduction

The dramatic effects of conventional exponential discounting on present values of costs and benefits that accrue in the distant future along with the issues of intergenerational equity that arise are well documented (see e.g. Portney and Weyant 1999, Pearce *et al.* 2003). The emergence of a long-term policy arena containing issues as diverse as climate change, nuclear build and decommission, biodiversity conservation, groundwater pollution, and the use of social Cost Benefit Analysis (CBA) to guide decision-makers in this arena has brought the discussion of long-run discounting to the fore. Discount rates that decline with the time horizon (Declining Discount Rates or DDRs) have often been touted as an appropriate resolution to what Pigou (1932) described as the 'defective telescopic faculty' of conventional discounting, and there has been much discussion about the moral and theoretical justification for such a strategy (see e.g. Dybvig *et al.* 1996, Sozou 1998, Weitzman 1998, 2001, Portney and Weyant 1999, Gollier 2002a). Of particular interest are the declining yet socially efficient discount rates resulting from the analysis of Weitzman (1998, 2004) and Gollier (2002a, 2002b, 2004) both of which appear to offer a theoretical path through the 'dark jungles of the second best' (Baumol 1968) and the intergenerational equity-efficiency trade-off contained therein.

If these theoretical solutions offer even a partial resolution of the problems of conventional discounting then it is clearly important that they can be operationalised and a schedule of DDRs can be determined. In the case of Gollier (2002a) and Weitzman (1998) it is uncertainty that drives DDRs, with regard to future growth of consumption and the discount rate respectively, thus the question of implementation is one of characterising the uncertainty of these primals in some coherent way. However, of these two approaches it is Weitzman (1998) that has proven to be more amenable to implementation mainly because the informational requirements stop at the characterisation of uncertainty, and do not extend to specific attributes of future generations' risk preferences as would be unavoidable in the case of Gollier (2002a, 2002b).[1]

[1]Weitzman (1998) assumes risk neutral agents for exposition, but this represents a special case of his general point. For realistic scenarios, determination of DDRs a la Gollier (2002a, 2002b) requires knowledge of the 4th and 5th derivatives of utility functions, something that he admits is very far from

Weitzman's Certainty Equivalent Discount Rate (CER) is derived from the expected discount factor and is therefore a summary statistic of the distribution of the discount rate. The level and behaviour over time of this statistic is clearly dependent upon the manner in which uncertainty is characterised and the two applications that exist have taken different approaches stemming from different interpretations of uncertainty. Weitzman (2001) defines uncertainty by the current lack of consensus on the appropriate discount rate for the very long term. His survey of professional economists results in a Gamma probability distribution for the discount rate which leads to the so-called 'Gamma discounting' approach, a version of which can also be seen in Sozou (1998). Apart from uncertainty his model has persistence in-built, the assumption being that each individual discounts the future at their preferred constant rate, that is each of the responses that make up the probability distribution remain constant over time.

More recently, Newell and Pizer (2003) (N&P, henceforth) suggest that while we are relatively certain about the current level of discount rates, there is considerable uncertainty in future. From this standpoint they assume that the past is informative about the future and characterise interest rate uncertainty by the parameter uncertainty typically found in any econometric model. They choose to describe the behaviour of the US long-term real interest rate with a reduced-form model. Their model is the direct analogue of the Vasicek (1977) model for the term structure of interest rates in the sense that only the conditional mean equation is specified and the conditional variance is held constant. In this respect, the authors get a working definition of the CER based upon an econometric model and estimation of the CER schedule comes from a forecasting simulation. Weitzman (2004) goes one step further and builds a "statistical optimal growth model" by combining a neoclassical economic model of optimal growth under uncertainty with a fully integrated Bayesian statistical model of estimating, updating and predicting the outcome of this uncertainty. His model is able to produce persistent uncertainty in the interest rate and as a result DDRs stemming mainly from the uncertainty over future technological progress. From a different point of view, mainly driven by the existing finance literature

being accomplished.

on the term structure of interest rates, Gollier (2004) reaches similar conclusions. He, specifically, finds that a positively correlated growth process leads to a decreasing yield curve in the case of a prudent representative agent due to increased uncertainty for the distant future. He also links his model with second order stochastic correlation and as a result to the Cox, Ingersoll and Ross model (1985) (CIR, henceforth) of the finance literature, introducing the analogue of heteroscedasticity in his process for the interest rate. In two simulation experiments, one including discrete jumps in the growth of consumption and the other parameter uncertainty, he provides evidence of DDRs and suggests that the discount rate should be as low as 1% for periods exceeding 400 years.

The aforementioned studies bring to light some interesting issues concerning the characterisation of the future path of interest rates. It is mainly persistence combined with uncertainty that leads to decline in discount rates over time. In the theoretical studies of Gollier and Weitzman, persistence is generated by the economy itself, while in N&P, the existence of persistence is an empirical question and it is the degree of persistence in the series that determines the rate of decline of the CER. In particular, N&P specify a simple AR(p) model of interest rate uncertainty, which limits the characterisation of uncertainty to a process in which the distribution of the permanent and temporary stochastic components is constant for all time. Such a process guarantees declining CERs, but it takes into account only the evolution of the mean of the process. As already mentioned their model is a discrete time version of the Vasicek (1977) continuous-time model in which the drift of the process is linear and mean-reverting, while the diffusion function is held constant. Since the seminal contribution of Vasicek (1977), an immense literature on the term structure of interest rates has produced interesting insights as to what drives efficient discount rates. The basic extensions mainly come from the specification of the variance of the process, namely the diffusion function. For example, CIR model the diffusion function as a linear function of the level of the interest rate, while Chan *et al.* (1992) allow the diffusion function to be any power function of the level of the interest rate. However, the aforementioned one-factor models display time-homogeneity, i.e. their parameters remain constant over time. It is reasonable to expect that the instantaneous return and volatility slowly evolve over time. In this respect, various efforts have been made to produce time-dependent models, such those of Ho and Lee (1986), Black *et al.* (1990), Hull and White (1990) and Black and Karasinski (1991). These models specify both the drift and the diffusion process of the instantaneous stochastic rate via time-varying functions of the level of interest rates.

The empirical issues stemming from the environmental literature on declining discount rates along with the development of an econometric model, versatile enough to reproduce the empirical regularities typically encountered in interest rate data are the main concern of this paper and we build upon the following points. Firstly, it is clear that if we believe that the past is informative about the future, it is important to characterise the past as accurately as possible. Indeed, the selection of the econometric model is of considerable moment in operationalising a theory of DDRs that depends upon uncertainty and defines the CER in statistical terms. Each specification differs in the assumptions made concerning the time series process, hence the forecasts of the interest rate and the attributes of the resulting schedule of the CER will differ accordingly. Secondly, the prescription of CBA will differ markedly depending upon the empirical schedule of discount rates employed, particularly for projects with a long time horizon such as climate change prevention. Moreover, model selection is also an empirical question. Typical misspecification testing and comparisons among various econometric models based upon their out-of-sample forecasting performance should guide model selection for the practitioner.

We revisit these issues for US interest rate data and show that misspecification testing generates a natural progression away from the simple AR(p) specification towards models which account for second-order dependence and explicitly consider changes in the time series process over time. We employ, for comparison purposes, the same data set of the US interest rates with N&P and show the policy implications of interest rate uncertainty and model selection in the value of carbon damages or sequestration.

The paper is organised as follows. In Section 2, we introduce the theory of the CER offered by Weitzman (1998), our methodology for model selection and the econometric models employed to replicate the stochastic nature of US interest rates. The results of the estimation and the simulations are presented in Section 3. Section 4 draws policy implications for model selection in the case of the value of carbon mitigation and Section 5 concludes the paper.

2 From Theory to Practice

2.1 The Certainty Equivalent Discount Factor and Rate

Discounting future consequences in period t back to the present is typically calculated using the discount factor P_t, where $P_t = \exp(-\sum_{i=1}^{t} r_i)$. When r is stochastic, the expected discounted value of a dollar delivered after t years is:

$$E(P_t) = E\left(\exp(-\sum_{i=1}^{t} r_i)\right) \tag{1}$$

Following Weitzman (1998) we define (1) as the *certainty equivalent discount factor*, and the corresponding *certainty-equivalent forward rate* for discounting between adjacent periods at time t as equal to the rate of change of the expected discount factor:

$$\frac{E(P_t)}{E(P_{t+1})} - 1 = \widetilde{r}_t \tag{2}$$

where $\widetilde{r}_t$ is the forward rate from period t to period $t+1$ at time t in the future, or the marginal discount rate. Gollier (2002a) shows that the certainty equivalent rate is the socially efficient discount rate in a risk neutral world – risk neutral agents are only concerned with the expected value of the discount factor rather than higher order moments – by showing that an arbitrage exists if this is not the case.[2] In effect this represents the economic theory underlying Weitzman's definition, however the behaviour of $\tilde{r}_t$ over time is dependent upon the nature of the uncertainty surrounding the discount rate. Weitzman (1998) and N&P show that $\widetilde{r}_t$ as defined in (2) is a declining function of time provided that there is sufficient persistence in the series over time.[3] This makes it clear

[2]Strictly, Gollier deals with the average certainty equivalent rate, however the same arguments hold as $t \to \infty$. His proof follows Dybvig *et al.* (1996).

[3]Weitzman (1998) gives a proof for a general but time invariant distribution function of $\widetilde{r}_t$. Weitzman (2001) estimates this distribution empirically as a Gamma distribution. Pearce *et al.* (2003) provide a

that operationalising this theory is an empirical question, requiring the determination of the stochastic nature of $\widetilde{r}_t$.

2.2 Parameterisation of Real Interest Rates

N&P employed a simulation method to forecast discount rates in the distant future, which was properly designed to account for uncertainty in the future path of interest rates and was mainly based on the estimation results of two econometric models, namely an autoregressive Mean-Reverting (MR) model and a Random Walk (RW) model. They estimated the following $AR(p)$ model for r_t:

$$\begin{aligned} r_t &= \eta + e_t \\ e_t &= \sum_{i=1}^{p} a_i e_{t-i} + \xi_t \end{aligned} \tag{3}$$

where $\xi_t \sim N(0, \sigma_\xi^2), \eta \sim N\left(\overline{\eta}, \sigma_\eta^2\right)$ and $\sum_{i=1}^{p} a_i < 1$ for the MR model, while $\sum_{i=1}^{p} a_i = 1$ for the RW model. The authors prove that in the case of an AR(1) model, the CER takes the following form:

$$\widetilde{r}_t = \overline{\eta} - t\sigma_\eta^2 - \sigma_\xi^2 f(\rho, t) \tag{4}$$

where $\overline{\eta}$ is the unconditional mean discount rate, ρ is the autoregressive coefficient, $f(\rho, t) = \frac{1-\rho^2-2\log(\rho)\rho^{t+1}\left(1+\rho-\rho^{t+1}\right)}{2(1-\rho)^3(1+\rho)}$ for MR and $f(\rho, t) = \frac{1}{12}(1 + 6t + 6t^2)$ for RW. It is straightforward to see that (4) is a declining function of t (See N&P for details).

This model, although simple, is successful in capturing the basic features of the underlying Data Generation Process (DGP) which lead to DDRs, namely persistence and uncertainty. However, given the abundance of models already designed to capture the dynamics of the interest rate data either in discrete or continuous time, it is hard to believe that simply modelling the mean of such a process is an adequate parameterisation of reality. As early as 1985, CIR introduce second-order dependence in the stochastic process of the interest rate by letting the conditional variance vary with the level of the

numerical example of the decline of the certainty equivalent discount rate for a uniform distribution.

interest rate.[4] The simpler discretised diffusion model motivated by the CIR model is the GARCH (1,1) model, in which the conditional variance depends on its own lag as well as the lag of squared innovations. However, when fitting a GARCH model to interest rates, one often finds that the parameter estimates imply that the conditional variance process is either integrated or explosive. Engle *et al.* (1987, 1990), Hong (1988), Harvey (1993) and Kees *et al.* (1997) document such a behaviour mainly for the US short term interest rates. In such cases, proper statistical testing usually cannot reject the hypothesis that the conditional variance of the process follows an integrated GARCH process (IGARCH). In our study, we employ the $AR(p)$ - $GARCH(l,m)$ model to account for both mean and volatility effects in the US interest rate process. Specifically our model is as follows:

$$\begin{aligned} r_t &= \eta + e_t \\ e_t &= \sum_{i=1}^{p} a_i e_{t-i} + \xi_t \\ \xi_t &= h_t^{1/2} z_t \\ h_t &= c + \sum_{i=1}^{m} \beta_i \xi_{t-i}^2 + \sum_{i=1}^{l} \gamma_i h_{t-i} \end{aligned} \tag{5}$$

where h_t is the conditional volatility of ξ_t (given all available information at time $t-1$) and $z_t \sim IIDN(0,1)$. In the case that $\sum_{i=1}^{m} \beta_i + \sum_{i=1}^{l} \gamma_i = 1$, we have an $AR(p)$ - $IGARCH(l,m)$ model.

Both the $AR(p)$ and $AR(p)$ - $GARCH(l,m)$ models assume that the parameters driving the stochastic process are constant over the sample period, i.e. they are time-homogenous. This is likely to be an unrealistic assumption for a period of 200 years and certainly for forecasting the CER over the long-term policy horizon in hand which, following N&P, extends for 400 years. It is well known that the behaviour of interest rates is strongly affected by the economic cycles as well as shocks destabilising them, i.e. periods of economic crisis. For example, in the US, during the period 1979 through 1982, the Federal Reserve Bank (FED) stopped its usual practice of targeting interest rates and decided to use non-borrowed reserves as a target instrument for monetary policy. As a

[4] Chan *et al.* (1992) extend the CIR model to include any power function for the diffusion function.

result, the volatility of US interest rates increased dramatically during that period. Other periods of high volatility of the US interest rates were the OPEC oil crisis (1973-1975), the October 1987 stock market crash and wars involving the US. Such turbulent periods are likely to induce persistence in volatility, which is often an artifact of the changes in the economic mechanism generating the interest rate (see Gray 1996). Lamourex and Lastrapes (1990) show that any structural shift in the unconditional variance is likely to lead to unreliable estimates of the GARCH parameters such that they imply too much persistence in volatility. In this sense, regime shifts are mistaken for periods of volatility clustering. Consequently, studies in the term structure literature have modelled discrete regime shifts in the spot interest rate process (Hamilton 1988, Das 1994, Gray 1996 and Naik and Lee 1997). These models typically posit a spot interest rate process that can shift randomly between two or more regimes (for example a low-mean and a high-mean regime). The diffusion and drift functions are kept the same but the specific parameter values are different in each regime. This makes the process time-heterogeneous. Each regime incorporates a different speed of mean-reversion to a different long-run mean and a different unconditional variance. Specifically, in our study we consider the following Regime-Switching (RS) model with two states:

$$\begin{aligned} r_t &= \eta_k + e_t \\ e_t &= \sum_{i=1}^{p} a_i^k e_{t-i} + \xi_t \end{aligned} \tag{6}$$

where $\xi_t \sim IIDN(0, \sigma_k^2)$, $k = 1, 2$ for the first and second regime, respectively. At any particular point in time there is uncertainty as to which regime we are in. The probability of being in each regime at time t is specified as a Markov 1 process, i.e. it depends only on the regime at time $t-1$. We define the probability that the process remains at the first regime as P, while the probability that the process remains at the second regime is Q. The matrix of the transition probabilities is assumed to be constant.[5]

[5] We define the following matrix of transition probabilities:

$$\begin{aligned} \Pr\text{ob}(R_t &= 1 \mid R_{t-1} = 1) = P, \ \Pr\text{ob}(R_t = 2 \mid R_{t-1} = 2) = Q \\ \Pr\text{ob}(R_t &= 2 \mid R_{t-1} = 1) = 1 - P, \ \Pr\text{ob}(R_t = 1 \mid R_{t-1} = 2) = 1 - Q \end{aligned}$$

The parameterisation of an RS model allows us to define a finite number of states that the economy goes through, which consequently affect the interest rate. However, it does not allow for cases that both the level and the variance of the process slowly evolve over time. Such an evolution can be captured by models with time-dependent parameters. In the continuous time literature, various models have been proposed in an effort to capture this time-dependence of parameters. These include the models of Ho and Lee (1986), Black *et al.* (1990), Hull and White (1990) and Black and Karasinski (1991). Fan *et al.* (2003) compare various specifications of both time-dependent and time-independent models and propose a time-varying coefficient model which captures better the time-variation of short-term dynamics of the interest rate. This finding, along with a similar conclusion of Ait-Sahalia (1996) who finds strong non-linearity of the drift for the US interest rate, leads us to introduce a time varying parameter model. We model the interest rate as a State Space (SS) process. More in detail, we specify an $AR(1)$ process with an $AR(p)$ coefficient as follows:

$$\begin{aligned} r_t &= \eta + \alpha_t r_{t-1} + e_t \\ \alpha_t &= \sum_{i=1}^{p} \eta_i \alpha_{t-i} + u_t \end{aligned} \tag{7}$$

where e_t and u_t are serially independent, zero-mean normal disturbances such that:

$$\begin{pmatrix} e_t \\ u_t \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_e^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix} \right). \tag{8}$$

This specification is able to capture non-linearities in the mean of the interest rate and accommodates changes in the conditional variance of the series under consideration. Tsay (1987) shows that the ARCH models can be regarded as special cases of Random Coefficient Autoregressive models (RCA), which are nested in the class of the $AR(1)$ model with an $AR(p)$ coefficient. A simple RCA model allows for the conditional variance to evolve with previous observations, accommodating in this way the high volatility observed

where R_t refers to the regime at time t.

in periods of high interest rates. With the addition of an $AR(p)$ structure to the coefficient of our model, we are able to capture both the volatility dynamics and the observed non-linearity in the drift of the interest rate process. This time-varying coefficient model can be thought of as an infinite regime-switching model which allows for a rather elevated degree of time-heterogeneity compared with the previous models.

Given the abundance of econometric models, our aim is to select the model that captures the dynamics of the data generating process in order to achieve an adequate description of the series under scrutiny. The complexity of the model and the restrictions it imposes should correspond to the level of uncertainty of the true data generating process. Otherwise, inference can be misleading and the forecasting performance of the model may be very poor. Common misspecification tests, such as tests for stationarity, autocorrelation, heteroscedasticity or parameter instability, will provide a benchmark to our selection procedure in conjunction with an out-of-sample forecasting exercise.

3 Empirical Results

3.1 Data

We use the US data employed by N&P for comparison purposes. More specifically, we use annual US market interest rates for long-term government bonds for the period 1798 to 1999. Starting in 1955, the nominal interest rates are converted to real interest rates by subtracting a ten-year moving average of the expected inflation rate of the CPI, as measured by the Livingston Survey of professional economists. For the previous years, expected inflation is assumed to equal zero and thus nominal and real interest rates coincide. The real interest rates are then converted to their continuously compounded equivalents. Finally, the estimation is based on a three-year moving average of the real interest rates series to smooth any short-term fluctuations, since we focus on the long-term behaviour of the series.[6] Following N&P, we estimate our models based on the logarithms of the series. This logarithmic transformation precludes negative rates and makes interest

[6] More details about the data can be found in N&P.

rate volatility more sensitive to the level of interest rates.[7]

3.2 Results

First of all, we test the stationarity of the US real interest rates. The results of a variety of unit-root tests are reported in Table A.1 of Appendix A.[8] These results generally favour the existence of a unit-root in the series, in line with the results of N&P. However, it is well-known that unit-root tests often lack the power to reject a false hypothesis of a unit-root for alternatives that lie in the neighbourhood of unity. Furthermore, mean shifts and non-linearities are often mistaken for unit-root behaviour (see, for example, Perron 1990 and Nelson *et al.* 2001). More importantly, it is difficult to believe that real interest rates become potentially unbounded with no economic forces at work to bring them back to some equilibrium, especially with two centuries of data. Albeit, for completeness, we estimate both a Random Walk (RW) and a Mean-Reverting (MR) model. Three lags are included in both models ($p = 3$).[9] Our estimates are identical to N&P and we do not discuss them extensively, for brevity. The MR model suggests conversion to a long-run mean of 3.69% at a very low speed though, as the sum of the autoregressive coefficients is as high as 0.976. Furthermore, tests for serial correlation in the residuals of the regression model suggest that mean dependence is sufficiently captured by this $AR(3)$ model. Not surprisingly though, this constant-variance model does a poor job in modelling the conditional volatility of interest rates as there is remaining autocorrelation in the squared residuals. Specifically, the Lagrange Multiplier (LM) test for autoregressive conditional heteroscedasticity (ARCH) in the residuals rejects the null hypothesis of homoscedasticity. In this respect, we estimate an $AR(3) - GARCH(1,1)$ model. In line with other empirical studies employing GARCH models to estimate the volatility of interest rates, we find that $\beta_1 + \gamma_1 = 1.007$, implying that the unconditional

[7]See N&P, footnote 15, pp.60 for a detailed discussion on this issue.

[8]We use the following unit root tests: the Augmented Dickey-Fuller test (Dickey and Fuller 1979), the Dickey-Fuller test with GLS detrending (Elliott *et al.* 1996), the Elliott-Rothenberg-Stock Point Optimal test (Elliott *et al.* 1996), the Phillips-Perron test (Phillips and Perron 1988), the KPSS test (Kwiatkowski *et al.* 1992) and the Ng-Perron test (Ng and Perron 2001).

[9]Throughout this paper, we use the Schwarz Information Criterion (SIC) to select the lag-length of the alternative models.

variance of the process is unbounded.[10] However, statistical tests indicate that β_1 and γ_1 sum up to unity, implying that the process of the conditional variance of the interest rate follows an integrated GARCH process. In this respect, we estimate an $AR(3) - IGARCH(1,1)$ model. The estimation results are reported in Table A.2 (Appendix A, Panel A). The estimates for the conditional mean remain the same in this setting, while the estimates for the conditional variance indicate that any shock is persistent in the sense that it remains important for future forecasts of all horizons.

However, as discussed above, this strong persistence in the volatility of the estimated GARCH model is an indication of a regime-switching mechanism in the generating process of the interest rate. In this mode, we estimate a two-regime model, where each regime is an AR(2) process. Table A.2 (Appendix A, Panel B) reports the estimates of this model. Both regimes are fairly persistent as indicated by the probabilities P and Q of the transition matrix which approach or even exceed 0.9. However, these regimes are distinct, as they display different characteristics. The first regime can be characterised as a "low-mean" regime, while the second as a "high-mean" one. The unconditional means for the two regimes are 3.28% and 5.55%, respectively. Different degrees of mean reversion are implied by the two regimes, as well. The "low-mean" regime mean-reverts quicker than the "high-mean" one as indicated by the sum of the autoregressive coefficients. The respective figures are 0.929 and 0.987, implying that our process is stationary in each regime. Moreover, the estimated transition matrix in combination with the estimated coefficients satisfy the condition for global second-order stationarity of the process, which is a desirable property as far as modelling the real interest rate is concerned.[11] Since such a type of model can just draw probabilistic assumptions about the state of the interest rate we are in, our estimates suggest that the probability (unconditional) of being in the "low-mean" regime is more than double the probability of being in the "high-mean" one (68% as opposed to 32%).[12] As a result, the estimated duration of the regimes is 7.5 years and 12 years for the low-mean and the high-mean regime, respectively. Furthermore, the

[10] Engle *et al.* (1990) report $\beta_1 + \gamma_1 = 1.0096$ for a portfolio of US securities, Kees *et al.* (1997) report $\beta_1 + \gamma_1 = 1.10$ for the one-month T-bills and Hong (1988) reports $\beta_1 + \gamma_1 = 1.073$.

[11] See Francq and Zakoian (2001) for the stationarity conditions.

[12] See Figure 1 for the estimated states over time.

first regime is more volatile than the second as indicated by the higher variance of the error term. Specifically, the estimated variance of the "low-mean" regime is 10 times greater than the variance of the "high-mean" one. This finding along with the estimated duration of the regimes leads us to assume that these regimes incorporate a business cycle effect over this 200-year period. As a result, periods with low real interest rates correspond to periods of slow growth or high inflation inducing uncertainty to the overall economy, while periods of high real interest rates correspond to periods of high growth and consequently confidence about the future state of the economy.

This business cycle effect or, more generally, the evolution of economic fundamentals might not be abrupt, switching from one state to the other. A gradual change in the evolution of the economy and interest rates as well might be captured better with a state space model. We specifically model the interest rate process as an $AR(1)$ process with an $AR(1)$ coefficient. The parameter estimates for this model are presented in Table A.2 (Appendix A, Panel C). The constant in our model suggests a minimum for the real interest rate, rather than a mean value, which is estimated at 1.67%. Furthermore, the autoregressive coefficient is strongly persistent.[13] This finding cannot in itself suggest any degree of mean reversion for the process as a whole, since the degree of mean reversion of the process changes over time. At the end of our sample the process of the interest rate displays a relatively quick mean reversion as suggested by a value of 0.47 of the relevant coefficient. Figure 2 in Appendix B shows the states of the estimated coefficient over time.

[13] Stability conditions for this process have been derived by Weiss (1985). Specifically, for a univariate AR(1) process with an AR(1) coefficient, i.e.

$$\begin{aligned} x_t &= \mu + \rho_t x_{t-1} + e_t, \\ \rho_t &= \phi \rho_{t-1} + v_t,\ Var(v_t) = q \end{aligned}$$

Weiss (1985) provides the following condition:

$$\begin{aligned} R + S^2(\infty) \quad : \quad &= \mu^2 + \frac{q}{1-\phi^2}(1 + 4\mu^2 + 8\mu^2 \lim_{n\to\infty} \sum_{j=1}^{n-1} \frac{n-j}{n} \phi^j) + \\ & \frac{2q^2}{(1-\phi^2)^2}(1 + \lim_{n\to\infty} \sum_{j=1}^{n-1} \frac{n-j}{n} \phi^{2j}) < 1 \end{aligned}$$

This condition is satisfied for our process.

3.3 Certainty-equivalent Discount Rates and Discount Factors

We follow N&P and simulate 100.000 possible future discount rate paths for each model starting in 2000 and extending 400 years into the future. For each model presented and estimated in the previous section the simulations are based on the estimates presented in Table $A.2$ (Appendix A, Panels A to C).[14] The initial value of the real interest rate is set at 4%, which as N&P argue reflects the best comparison with a constant rate. In Appendix C, we briefly describe the simulation method for each estimated model. We then calculate the certainty-equivalent discount rate employing equation (2). The simulated expected discount factors and the corresponding CERs are reported in Tables 1 and 2, respectively for the various models into consideration.

The first column of Table 1 displays the discount factors based on a constant 4% rate with the remaining columns corresponding to the estimated models. As expected, the models produce considerably different discount factors and the differences between them are evident even from the first 60 years. For example, for a 60 year horizon the SS model produces substantially higher valuations than the rest of the models (the difference is over 50 % in some cases). Overall, the higher valuations come from either the SS or the RW model. The present value of $1 delivered after 100 years is $0.05 and $0.08 according to RW and SS respectively. The corresponding value for the rest of the models is about $0.02. At the end of the period under examination, the RW model is the one that retains the higher value followed by the AR-IGARCH and the SS models.

{INSERT TABLE 1 HERE: 1: Discount Factors}

Naturally, the differences among discount factor projections relevant to each model are reflected in the projected schedule of the CERs. All the models accommodate declining interest rates mainly stemming from the persistence and uncertainty built in them. They differ, however, at the path they follow and the terminal values they attain. For example, SS and RW produce the lower rates for the first 100 years, reaching a CER of around 2% (half the initial value). During the same period, the MR and the AR-IGARCH models follow similar paths yielding a reduction of just 50 basis points. In the case of RS, the

[14]The reader is referred to N&P for the estimates of the RW and the MR models (Table 1, page 63).

CER increases slightly due to some overshooting during the first 40 years. Except for this overshooting, the RS model regains its quick declining path for the rest of the period reaching a rate of 0.7% after 400 years. The highest terminal rate is produced by the SS model, which projects a rate of 1.6%, followed by MR at 1.4%.

{INSERT TABLE 2 HERE: 2: CERs}

In summary, the forecasts of the alternative models differ substantially. In this respect, we need to evaluate the models with respect to their predictive ability. Typical misspecification testing has shown that a constant coefficient model may not be able to fully capture the dynamics of the US interest rates over the period examined. Along this line of reasoning, we suggested two time-varying coefficient models (RS and SS), one accommodating abrupt changes and the other allowing for a gradual change over time in the generating mechanism of the interest rates. These two models seem eminently preferable to the constant coefficient models. In the following subsection, we perform an out-of-sample forecast exercise to select among the various models.

3.4 Model Selection

Evaluating the out-of-sample forecasting performance of the models under consideration for the long run is impossible due to limitation of data, as forward rates exist for a maximum period of 30 years. However, we attempt to discriminate between these models on the grounds of their forecasting performance over a 30-year horizon using available real data. We specifically make use of annual forward rates suggested by the term structure of the inflation-indexed US government bonds. Then, we calculate the commonly-used Mean Square Forecast Error (MSFE) and judge the models by this criterion. Alternatively, we calculate four modified MSFE criteria by incorporating four kernels[15] which weigh observations by their relevant proximity to the present. The results are presented in Table 3.

{INSERT TABLE 3: Average MSFEs}

Interestingly, the various specifications of the MSFE criterion unanimously rank the

[15] The Bartlett(B), the Parzen(P), the Quadratic-Spectral (QS) and the Tukey-Hanning (TK) kernels are the weighting functions used in our evaluation.

SS model first followed by the RS model in most of the cases. The AR-IGARCH model ranks third followed by MR and then RW.

In sum, if we select a model on the basis of its ability to characterise the past and its accuracy concerning forecasts of the future, we are inclined to accept the SS model as the best model (among the estimated models) to describe the US real interest rates. Our second best choice would be the RS model.

4 Policy Implications of Model Selection

The foregoing has established the importance of model selection in determining a schedule of declining discount rates for use in CBA. The differences that arise from alternative specifications of the time series process have been revealed and a method for selecting one model over another has been proposed. In this section we highlight the policy implications of declining discount rates and the impact of model misspecification by considering the same case study as N&P, that is, climate change and the value of carbon sequestration.[16] We establish the present value of the removal of 1 ton of carbon from the atmosphere, and hence the present value of the benefits of the avoidance of climate change damages for each of the specified models. To understand what follows it is important to be familiar with the profile of benefits resulting from the removal of 1 ton of carbon from the atmosphere. We use the estimates taken from the DICE model of Nordhaus and Boyer (2000) shown in Figure 3 (Appendix B). Table 4 shows the present value per ton of carbon emissions when evaluated using the schedule of discount rates associated with each of the models described in Section 3.2.

{INSERT TABLE 4 HERE}

The RS model gives the lower valuations followed by the conventional 4% discounting. Interestingly, the SS model gives the higher valuation followed by the RW model. For example, the present value of carbon emissions reduction is over 150 % larger in the case of the SS model compared to the case of constant discounting at 4 %. On the other hand, the present value of the removal of 1 ton of carbon emissions from the atmosphere increases

[16] See N&P for the assumptions concerning the modeling of carbon emissions damages.

by only 12 % based on the MR's forecasts compared to the constant rate discounting approach.

The preceding discussion has argued that the RS and SS models are to be preferred over the others since they allow for changes in the interest rate generating process and have desirable properties. From the policy perspective we have established that both these models provide well specified representations of the interest rate series. However, the RS model provides roughly equivalent values of carbon to the constant discounting rate values (there is a 9% difference), while the SS model produces values that are up to 150% higher than those of the constant rate.

The disparity between the RS and the SS models, and the proximity of the carbon values generated by the former to those generated by conventional constant discounting represents a clear signal of the policy relevance of model selection in determining the CER. It is crucial from a policy perspective to make a clear judgment as to which of the two models (RS and SS) is most appropriate to the case in hand. Our forecasting exercise reveals that the SS model is preferable to the RS model due to its lower MSFE for the 30-year horizon. Hence in the context of SS the carbon values are increased by 150% compared to conventional discounting and 40% compared to N&P's approach. In short, in the US context, the selection of econometric models on the basis of forecasting performance, and the preferred schedule of discount rates makes climate change prevention a more desirable investment.

5 Conclusions

In response to the need to appraise projects over very long time horizons, a number of theoretical discussions have arisen concerning the appropriateness of discount rates that fall with the time horizon considered. Such Declining Discount Rates (DDRs) would add greater weight to the costs and benefits that accrue to future generations and thereby at least partially address the issue of inter-generational equity that so often besets the long term policy arena.

Weitzman's (Weitzman 1998) theoretical justification for DDRs depends upon uncertainty of the discount rate and therefore the operationalising of this theory is highly dependent upon the manner in which one interprets and characterises uncertainty. Weitzman (2001) suggested that it was the lack of consensus about the correct discount rate to employ in the far distant future that was the source of uncertainty and his estimated Gamma distribution provided the means of operationalising this theory and determining the declining Certainty Equivalent Rate (CER). Newell and Pizer (2003) (N&P) took an alternative view, accounting for the uncertainty through an econometric forecasting approach.

This paper builds on N&P's approach in determining DDRs and it makes the following points concerning the model selection and the use of DDRs in general. Firstly, N&P's approach is predicated upon the assumption that the past is informative about the future and therefore characterizing uncertainty in the past can assist us in forecasting the future and determining the path of CERs. We have argued that if one subscribes to this view it is important to characterise the past as well as possible by correctly specifying the model of the time series process. This is particularly so when dealing with lengthy time horizons where the accuracy of forecasts is important. Indeed the selection of the econometric model is of considerable moment in operationalising a theory of DDRs that depends upon uncertainty, because econometric models contain different assumptions concerning the probability distribution of the object of interest. We have shown that when modelling the US interest rate data, the econometric model should allow for changes over time in the data generating process and that state space and regime switching models are likely to be appropriate.

Our estimations, simulations and case study bear out this assertion. The path of the CER differs considerably from one model to another and therefore each places a different weight upon the future. The policy implications of these estimates is revealed in the context of a case study that calculates the present value of carbon emissions reduction. The utilisation of a state space model to estimate the discount factors results in an increase of 150% in the present value of carbon emissions reduction compared to a constant rate discounting approach.

References

[1]Ait-Sahalia, Y. (1996). Testing Continuous-time Models of the Spot Interest Rate. *Review of Financial Studies* 9, 385-426.

[2]Baumol, W.J. (1968). On the Social Rate of Discount. *American Economic Review* 57, 788-802.

[3]Black, F., E. Derman and W. Toy (1990). A One-factor Model of Interest Rates and its Application to Treasury Bond Options. *Financial Analysts Journal*, 33-39.

[4]Black, F. and P. Karasinski (1991). Bond and Option Pricing when Short Rates are Lognormal. *Financial Analysts Journal*, 52-59.

[5]Chan, K.C., G.A. Karolyi, F.A. Longstaff and A.B. Sanders (1992). An Empirical Comparison of Alternative Models of the Short-term Interest Rate. *Journal of Finance* 47, 1209–1227.

[6]Cox, J.C., J.E. Ingersoll Jr. and S.A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica* 53, 385-407.

[7]Das, S.R. (1994). Mean Rate Shifts and Alternative Models of the Interest Rate. Theory and Evidence. *Working Paper, New York University.*

[8]Dickey, D.A. and W.A. Fuller. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* 74, 427–431.

[9]Dybvig, P.H., J.E. Ingersoll Jr. and S.A. Ross (1996). Long Forward and Zero-Coupon Rates Can Never Fall. *Journal of Business* 69, 1-24.

[10]Elliott, G., T.J. Rothenberg and J H. Stock (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64, 813-836.

[11]Engle, R.F., D.M. Lilien and R.P. Robins (1987). Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica* 55, 391-407.

[12]Engle, R.F., V.K. Ng and M. Rothchild (1990). Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills. *Journal of Econometrics* 45, 213-237.

[13]Fan, J., J. Jiang, C. Zhang and Z. Zhou (2003). Time-dependent Diffusion Models for Term Structure Dynamics and the Stock Price Volatility. *Statistica Sinica* 13, 965-992.

[14]Francq, C. and Zakoian J.-M. (2001). Stationarity of Multivariate Markov-switching ARMA Models. *Journal of Econometrics* 10, 339–364.

[15]Gollier, C. (2002a). Time Horizon and the Discount Rate. *Journal of Economic Theory* 107 (2), 463-473.

[16]Gollier, C. (2002b). Discounting an Uncertain Future. *Journal of Public Economics* 85, 149-166.

[17]Gollier, C. (2004). The Consumption-based Determinants of the Term Structure of Discount Rates. *Mimeo, University of Toulouse.*

[18]Gray, F.S. (1996). Modelling the Conditional Distribution of Interest Rates as a Regime-switching Process. *Journal of Financial Economics* 42, 27-62.

[19]Hamilton, J.D. (1988). Rational Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates. *Journal of Economic Dynamics and Control* 12, 385-423.

[20]Harvey, A.C. (1993). Long Memory in Stochastic Volatility. *Mimeo, London School of Economics.*

[21]Ho, T.S.Y. and S.B. Lee (1986). Term Structure Movements and Pricing Interest Rate Contingent Claims. *Journal of Finance* 41, 1011-1029

[22]Hong, C.H. (1988). The Integrated Generalised Autoregressive Conditional Heteroskedastic Model: The process, Estimation and Monte Carlo Experiments. *Working Paper, University of California*, San Diego, CA.

[23]Hull, J. and A. White (1990). Pricing Interest Rate Derivative Securities. *Review of Financial Studies* 3 (4), 573-592.

[24]Kees, G.K., F.G.J.A. Nissen, P.C. Schotman and C.C.P. Wolff (1997). The Dynamics of Short-term Interest Rate Volatility Reconsidered. *European Finance Review* 1, 105–130.

[25]Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure are we that Economic Time Series have a Unit Root? *Journal of Econometrics* 54, 159-178.

[26]Lamourex, G. and W. Lastrapes (1990). Persistence in Variance, Structural Change, and the GARCH Model. *Journal of Business and Economic Statistics* 23, 225–234.

[27]Naik, V. and M.H. Lee (1997). Yield Curve Dynamics with Discrete Shifts in Economic Regimes. Faculty of Commerce, *University of British Columbia, Working Paper.*

[28]Nelson, C.R, J. Piger and E. Zivot (2001). Markov Regime Switching and Unit-root Tests. *Journal of Business and Economic Statistics* 19 (4), 404-415.

[29]Newell, R. and W. Pizer (2003). Discounting the Benefits of Climate Change Mitigation: How Much do Uncertain Rates Increase Valuations? *Journal of Environmental Economics and Management* 46 (1), 52-71.

[30]Ng, S. and P. Perron (2001). Lag Length Selection and the Construction of Unit Root Tests with Good size and Power. *Econometrica* 69 (6), 1519-1554.

[31]Nordhaus, W.D. and J. Boyer (2000). Warming the World: Economic Models of Global Warming, MIT Press, Cambridge, MA

[32]Pearce, D., B. Groom, C. Hepburn and P. Koundouri (2003). Valuing the Future: Recent Advances in Social Discounting. *World Economics* 4 (2), 121-141.

[33]Perron, P. (1990). Testing for a Unit Root in a Time Series with a Changing Mean. *Journal of Business and Economic Statistics* 8 (2), 153-162.

[34]Phillips, P.C.B. and P. Perron (1988). Testing for a Unit Root in Time Series Regression. *Biometrika* 75, 335-346.

[35]Pigou, A. (1932). *The Economics of Welfare*, 4th edition, Mac Millan, London.

[36]Portney, P. and J. Weyant (eds) (1999). Discounting and Intergenerational Equity. Washington DC: Resources for the Future.

[37]Sozou, P.D. (1998). On Hyperbolic Discounting and Uncertain Hazard Rates. *Proceedings of the Royal Society of London Series B-Biological Sciences* 265 (1409), 2015-2020.

[38]Tsay, R. (1987). Conditional Heteroskedasticity Time Series Models. *Journal of the American Statistical Association 82,* 590-604.

[39]Vasicek, O. (1977). An Equilibrium characterisation of the Term Structure. *Journal of Financial Economics* 5, 177-188.

[40]Weiss, A. (1985). The Stability of the AR(1) Process with an AR(1) Coefficient. *Journal of Time Series Analysis* 6 (3), 181-186.

[41]Weitzman, M. (1998). Why the Far Distant Future Should be Discounted at its Lowest Possible Rate. *Journal of Environmental Economics and Management* 36, 201-208.

[42]Weitzman, M. (2001). Gamma Discounting. *American Economic Review* 91 (1), 261-271.

[43]Weitzman, M. (2004). Discounting a Distant Future whose Technology is Unknown. *Mimeo, Harvard University.*

Table 1. Certainty Equivalent Discount Factors

Model Year	4% Constant	Mean Reverting	Random Walk	AR IGARCH	Regime Switching	State Space
1	0.96154	0.96154	0.96154	0.96154	0.96154	0.96154
20	0.45639	0.45906	0.46177	0.45876	0.45390	0.56424
40	0.20829	0.21661	0.22917	0.21250	0.19576	0.33136
60	0.09506	0.10471	0.12480	0.10062	0.08458	0.20296
80	0.04338	0.05150	0.07777	0.04894	0.03700	0.12889
100	0.01980	0.02567	0.05082	0.02455	0.01647	0.08408
150	0.00279	0.00476	0.02333	0.00529	0.00238	0.03132
200	0.00039	0.00095	0.01830	0.00178	0.00041	0.01255
250	0.00006	0.00022	0.01119	0.00104	0.00010	0.00526
300	0.00001	0.00006	0.00890	0.00086	0.00003	0.00227
350	0.00000	0.00002	0.00715	0.00080	0.00002	0.00100
400	0.00000	0.00001	0.00669	0.00078	0.00001	0.00044

Table 2. Certainty Equivalent Discount Rates

Model Year	Mean Reverting	Random Walk	AR IGARCH	Regime Switching	State Space
1	4.00	4.00	4.00	4.00	4.00
20	3.91	3.85	3.96	4.22	2.79
40	3.76	3.46	3.88	4.31	2.59
60	3.65	3.08	3.74	4.26	2.38
80	3.58	2.60	3.60	4.18	2.23
100	3.51	2.17	3.42	4.09	2.10
150	3.36	1.39	2.75	3.79	1.91
200	3.16	0.94	1.62	3.31	1.79
250	2.87	0.75	0.65	2.46	1.72
300	2.43	0.56	0.23	1.83	1.67
350	1.87	0.43	0.09	0.95	1.64
400	1.41	0.34	0.04	0.70	1.61

Table 3. Average MSFEs

Model Criterion	Mean Reverting	Random Walk	AR IGARCH	Regime Switching	State Space
AMSFE	2.058	2.171	2.102	2.323	1.832
AMSFE (B)	1.692	1.724	1.692	1.687	1.499
AMSFE (P)	1.725	1.746	1.720	1.683	1.426
AMSFE (QS)	0.842	0.870	0.848	0.879	0.760
AMSFE (TH)	1.769	1.797	1.765	1.738	1.550

Notes: The weighting functions are as follows: Bartlett(B), Parzen(P), Quadratic-Spectral (QS) and Tukey-Hanning (TK).

Table 4. Value of Carbon Damages

Model	Carbon Values ($/tc)	Relative to Constant Rate	Relative to Mean Reverting	Relative to Random Walk
Regime-Switching	5.22	-9.0%	-18.8%	-49.4%
Constant (4.0%)	5.74	—	-10.7%	-44.4%
AR-IGARCH	6.37	11.0%	-0.9%	-38.3%
Mean Reverting	6.43	12.0%	—	-37.7%
Random Walk	10.32	79.8%	60.5%	—
State Space	14.44	151.6%	124.6%	39.9%

Appendix A: Tables

Table A.1: Unit Root Tests

Test	Lags /Bandwidth	t-stat.	5% critical value	Decision
ADF	13	-2.314	-2.877	non-stationary
Phillips-Perron	12	-2.016	-2.876	non-stationary
DF-GLS	13	-0.473	-1.942	stationary
ERS Point-Optimal	12	19.733	3.170	non-stationary
Ng-Perron	12	-0.824	-8.100	non-stationary
KPSS	15	1.158	0.463	non-stationary

Notes: SIC is employed to determine the lag-length of the series. The kernel sum-of-covariances estimator with Parzen weights is used, while the bandwidth is determined based on the Newey-West bandwidth selection method.

Table A.2: Estimation Results

Panel A: AR(3)-IGARCH(1,1) model			
Coefficient	Estimate	Std. Error	t-stat.
n	1.330	0.104	12.811
a_1	1.951	0.085	23.033
a_2	-1.322	0.156	-8.472
a_3	0.355	0.080	4.441
c	0.000	0.000	3.236
β_1	0.442	0.092	4.805
Panel B: Regime Switching model			
Coefficient	Estimate	Std. Error	t-stat.
n_1	1.189	0.128	9.327
a_1^1	1.589	0.078	20.36
a_2^1	-0.660	0.086	-7.630
n_2	1.714	0.238	7.206
a_1^2	1.787	0.050	35.55
a_2^2	-0.800	0.049	-16.395
σ_1^2	0.004	0.001	5.651
σ_2^2	0.000	0.000	6.070
P	0.867	0.058	14.934
Q	0.917	0.035	25.976
Panel C: State Space model			
Coefficient	Estimate	Std. Error	t-stat.
n	0.510	0.082	6.185
n_1	0.990	0.002	494.9
$ln(\sigma_e^2)$	-9.158	1.324	-6.917
$ln(\sigma_u^2)$	-6.730	0.144	-46.63

Appendix B: Figures

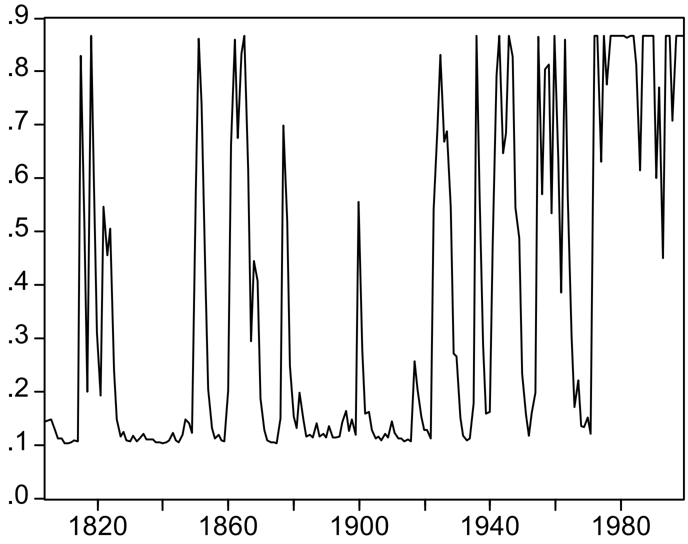

Figure 1: Filter Probabilities of the Regime Switching Model

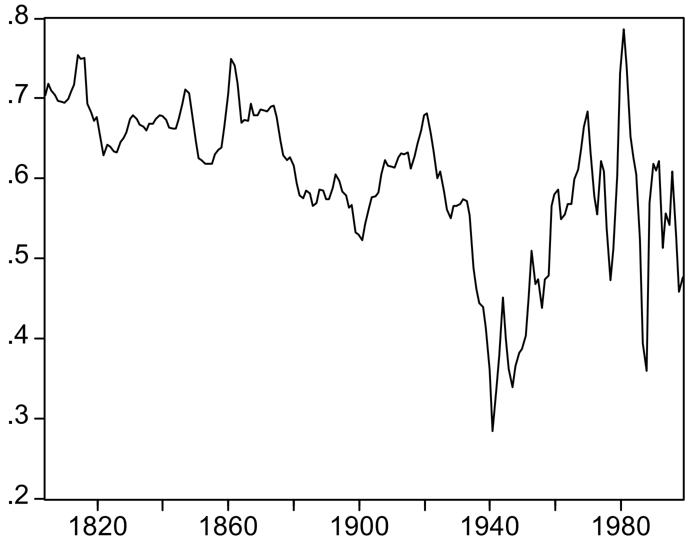

Figure 2: Evolution of the AR(1) Coefficient in the State-Space Model.

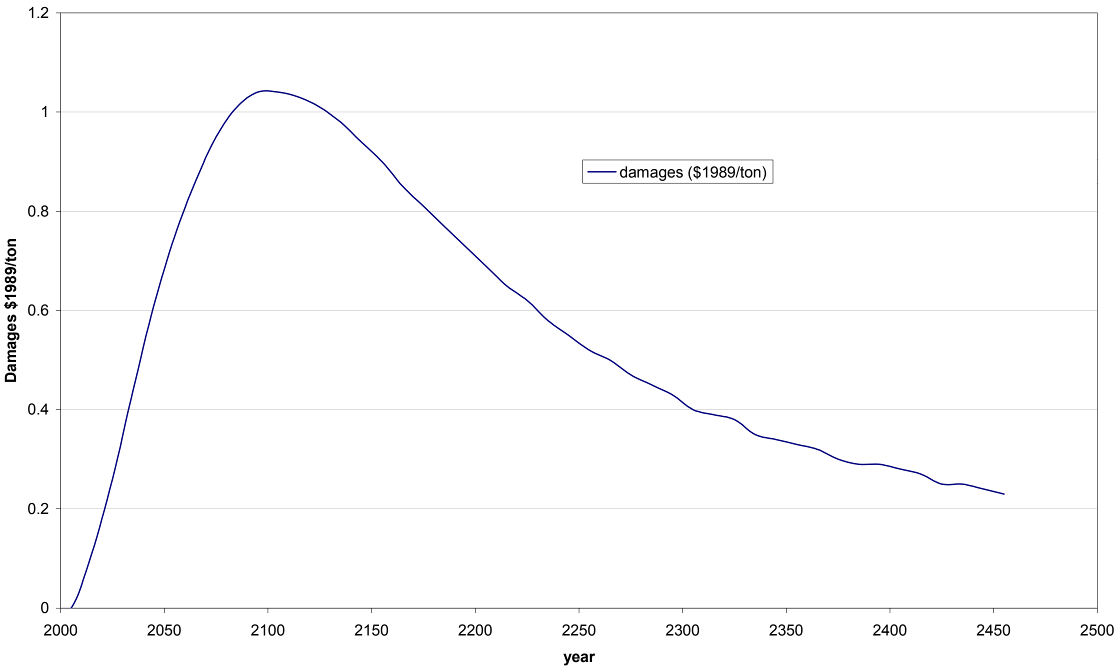

Figure 3: The Profile of Carbon Damages.

Appendix C. Simulation Methodology

Mean Reverting Model: We employ a multivariate normal distribution to draw random values for the coefficients of (3) taking into account the estimated variance-covariance matrix of the coefficients. Another draw from a normal distribution is employed for the estimated variance. Given this set of random parameters, we generate a future path of the interest rate. We repeat the same procedure to generate 100.000 random paths of the interest rates.

Random Walk Model: As previous.

AR(3)-IGARCH (1,1): The simulation methodology is similar to the MR model. However, in this case we use the multivariate normal distribution to obtain random draws for both the conditional mean and conditional variance parameters.

Regime Switching: The RS model offers the most computationally intensive simulation and is conducted as follows. First, we generate random values for the probabilities P and Q from a $Beta(k, j)$ distribution. The values of the parameters k and j of the Beta distribution are properly chosen in order to correspond to a Beta distribution with mean and standard deviation equal to the ones estimated. Specifically, in the case of P we set k and j equal to 28.8 and 4.42 respectively. The corresponding values for Q are 55.17 and 5, respectively. Using the random values of P and Q, we calculate the probability of being in each regime for each of the future 400 years, namely P_t and Q_t. A univariate normal distribution is used to get random draws for σ_1^2 and σ_2^2 separately according to the estimates presented in Table A.2 (Panel B). Similarly to our previous simulations, the random values for the coefficient estimates, n_1, n_2, $a^1_{1,}$ $a^1_{2,}$ a^2_1 and a^2_2 are drawn from a multivariate normal distribution. Then, we simulate the future interest rate path 100.000 times on the grounds of the probabilities P_t and Q_t and the random draws of the coefficients.

State Space: The simulation design for the SS model is straightforward as we randomly draw the coefficient values from univariate normal distributions according to the estimated values (Table A.2 (Panel C)). We then simulate the future path of interest rates in a similar way to the other models.