The Climate Change Learning Curve

Andrew J. Leach *

HEC Montréal, CIRANO, CIRPÉE

December 18, 2005

Abstract

The purpose of this paper is to confront economic models of climate change with the reality that limited information exists with which to form expectations about the evolution of the climate. A key element in the tension between those who believe we should impose aggressive climate change mitigation policies and those who do not is the question of whether we are merely in a long period of shock-induced, above average temperatures or if observed increases in temperature are a result of carbon emissions. This paper characterizes learning dynamics resulting from the use of observations of temperature to update beliefs about two key characteristics of global climate: the persistence of natural trends and the sensitivity of temperature to atmospheric carbon levels. This paper shows that, contrary to predictions in the literature that uncertainty may be resolved very quickly, the time to learn the true processes may be in the order of thousands of years. Further, this paper shows the effects of uncertainty on the likelihood that observations from the statistical record lead to important estimate and policy errors.

Key words: Climate Change; Bayesian Learning; Environmental Regulation; Growth; Pollution; Dynamic Programming; Precautionary Principle.

JEL classification: Q25; Q28; D83; D81; C61; C63; E1; E61; H4.

* I would like to thank the Editor, two anonymous referees, Robert Clark, David Kelly, Chris Ferrall, Bernard Sinclair-Desgagné, Ross McKitrick, and Michelle Reinsborough, as well as seminar participants at Université Laval and participants at the Canadian Resource and Environmental Economics Study Group, 2003, the Montreal Macroeconomics Reading Group, and the 2004 CIREQ Resource and Environmental Economics Workshop for helpful comments. The usual disclaimer applies.

Email address: andrew.leach@hec.ca (Andrew J. Leach).

Submitted to the Journal of Economic Dynamics and Control 18 December 2005

1 Introduction

A large and growing literature in economics addresses the challenge of developing optimal climate change policy in the face of uncertainty and expected future learning. Both the persistence of temperature changes (whether natural or anthropogenic) and the degree to which greenhouse gas (GHG) accumulation causes temperature change will be important for policy formation. In this paper, a model which captures the need to use a limited amount of information to form expectations about a complex system in order to set climate policy is developed. While there exist reasonable data describing the recent evolution of both temperature and atmospheric GHG accumulation, separately identifying the sources of temperature change as natural or anthropogenic solely based on the statistical record leads to significant uncertainty surrounding the relative magnitudes of these effects. If we assume knowledge of the natural process which governs temperature evolutions, then the process of identifying the effect of carbon is made to appear much less complex. This paper describes the nature of the uncertainty that exists over the mechanism of climate change through an empirical exercise, then characterizes the dynamics which are likely to arise as this uncertainty is resolved using a reduced form, learning experiment. Finally, learning and uncertainty are imposed in an optimal policy model to characterize how uncertainty is likely to affect policy choices and vice-versa.

The benchmark contributions to climate change economics are Manne and Richels (1992), Manne, Mendelsohn and Richels (1995), Nordhaus (1994), and Nordhaus and Boyer (2000). Each of these contain extensive reference to uncertainty, but generally treat uncertainty only through sensitivity analysis, reporting results for various parameter vectors. Pizer (1999) introduces a model where the regulator specifically accounts for parameter uncertainty in the social planning decision. Related papers on active learning to resolve uncertainty about the value of parameters governing climate change and the damages it may cause include Kolstad (1994, 1996, 1997), Ulph and Ulph (1997), Kelly, Kolstad, and Mitchell (1999), Kelly and Kolstad (1999), and Karp and Zhang (2004).

Kelly and Kolstad (1999) is closely related to the exercise undertaken here. This paper proposes a model in which a social planner uses information from temperature realizations to update prior beliefs about the temperature response to atmospheric GHG levels. The planner chooses the optimal level of savings and emissions control conditional on current knowledge of the mechanism of climate change at each point in time, updates these beliefs, and thus adjusts their actions, conditional on observations of climate data. Learning is Bayesian, thus the planner is using information in an optimal manner. A key result shown with the model is that the expected learning time (the time after

which parametric uncertainty is essentially removed from the planner's problem) is 90-160 years. The results also show that there is a tradeoff between the benefits of controlling emissions and information.

I first extend the results of Kelly and Kolstad (1999) by exploring, in a reduced-form environment, the dynamics of learning where uncertainty exists over the values of two parameters which jointly determine the evolution of global climate. While Kelly and Kolstad evaluates the expected time to resolve uncertainty about the effect of atmospheric GHG accumulation on temperature, I build on these results by examining the implications of adding uncertainty over the value of a parameter governing the persistence of temperature changes. Where uncertainty exists over these two parameter values simultaneously, results of a numerical experiment show that uncertainty is substantially more persistent, such that an increase in learning times of hundreds of years is possible. Further, it is shown that the inclusion of uncertainty in two dimensions greatly increases the likelihood that observations from the statistical record will support a mis-estimation of the true process of climate change.

The second contribution of this paper is to examine the effect of imposing the same type of uncertainty described above in an optimal policy model with learning. I use a modified Nordhaus and Boyer (2000) integrated assessment model (IAM), for which general results are already well known. The model is calibrated to economic data, and simulations are used to provide predictions on two effects. First, simulations are used to re-examine the dynamics of the resolution of uncertainty, and examine how these dynamics may be altered under differing assumptions about the true process. Second, the decisions of the planner characterize optimal carbon emissions control where beliefs have differing accuracy and levels of uncertainty relative to a certainty benchmark.

Simulations of the IAM show that the planner will under-regulate where uncertainty is present, which may decrease the persistence of uncertainty by increasing variance of emissions. It is also shown that the likelihood of persistent errors in estimates of the severity of climate change may lead to long-term inefficient policies. Sensitivity analysis shows that learning may be faster in expectation, but much more subject to errors, when the initial beliefs of the planner imply an under-estimate the severity of climate change.

This paper proceeds as follows. Section 2 presents the climate model, discusses the Bayesian learning approach, and presents a numerical learning experiment. Section 3 presents the climate and economy model with learning. Section 4 presents the solution algorithm. The model is calibrated in Section 5, and Section 6 presents simulations and optimal policy results. Section 7 concludes.

2 Learning about Two Causes of Climate Change

2.1 Model of Climate Change

The climate system is represented as in the DICE-99 model presented in Nordhaus and Boyer (2000); an autoregressive, distributed lag model where climate variables evolve as a function of GHG emissions inputs. The model, presented below for clarity of notation, is such that observed changes in temperature may reflect the persistence of natural, stochastic events or the effect of GHG emissions.

Let m_t (m_b) represent the current (pre-industrial) accumulation of carbon in the atmosphere.[1] Atmospheric carbon stock decays naturally at rate $\delta_m \in (0,1)$, and is augmented by emissions E_t according to the following law of motion:

$$m_{t+1} = E_t + (1-\delta_m)(m_t - m_b) + m_b. \tag{2.1}$$

Deviations from the mean of global temperature are modeled as a stochastic, first-order autoregressive process with drift generated by accumulated carbon in the atmosphere. Let G and O represent global surface and ocean temperature deviations respectively, and let the law of motion for G be given by:

$$G_{t+1} = \lambda_1 G_t + \omega O_t + \eta \left[\frac{\log\left(\frac{m_t}{m_b}\right)}{\log(2)} \right] + u_t,\ u \sim NID(0, \sigma_u^2). \tag{2.2}$$

Climate change is buffered in the short run by thermal inertia, captured through slowly changing ocean temperatures which are modeled as a deterministic, autoregressive process with parameter $\lambda_2 \in (0,1)$ as:

$$O_{t+1} = \lambda_2 O_t + (1-\lambda_2) G_t. \tag{2.3}$$

This stylized climate model allows us to parameterize an important variable of policy interest: the long-run temperature change associated with a doubling of atmospheric carbon, denoted by $G_{2C} = \frac{\eta}{1-\lambda_1-\omega}$. This expression is only informative where $|\lambda_1 + \omega| \in [0,1)$, a condition which assures the convergence of temperature to a long run equilibrium for any accumulation of carbon. All climate changes in the model are reversible so long as values of λ_1 and ω satisfy this condition.

[1] While this paper uses carbon mass, carbon may be measured in concentration (ppmv) or mass (GtC) of carbon or CO_2, as long as m_t, m_b, and the contribution of emissions to accumulated carbon are measured in the same units.

Schlesinger et al. (undated) explores in detail the use of statistical techniques to estimate parameter values for climate models of the class described in equations (2.1)-(2.3). I use similar techniques to investigate whether is it reasonable to assume that the stationarity condition described above holds, and to characterize uncertainty over values of G_{2C} and underlying parameters of the climate model. In particular, the parameters of the following restricted version of the model are estimated: [2]

$$G_{t+1} = \hat{\lambda} G_t + \hat{\eta} \left[\frac{\log\left(\frac{m_t}{m_b}\right)}{\log(2)} \right] + u_t,\ u \sim NID(0, \sigma_u^2). \tag{2.4}$$

Two sets of global temperature data are used; the Jones et al. (2005) data, which track temperature anomalies for years 1860-2000, and predictions generated by the Hadley Center HadCM-2 climate model (Johns et al., 1997) for years 2000-2100 under the IS-92a emissions scenario. [3] These data are matched with historic atmospheric carbon concentration data and predictions under the IS-92a emissions scenario from Joos and Siegenthaler (1999). [4] The estimation results are shown in the first and second columns of Table 1.

The first question of interest is the stationarity of temperature. From a policy point of view, stationarity has important implications since there is a strong discontinuity between the benefits to emissions abatement policy if temperatures are stationary versus if they are potentially growing without bound. Augmented Dickey-Fuller (ADF) tests for the presence of a unit root, allowing for a forcing trend, are performed using each of the data sets. The results of these tests are presented the third and fourth columns of Table 1. In both cases, the null hypothesis that temperature data exhibit a unit root is rejected. This provides support for the traditional assumption that the true law of motion for temperature is a stationary process.

[2] The restriction that $\lambda_2 = 0$ is imposed on the climate model. Under this assumption, the estimate of $\hat{\lambda}$ in equation (2.4) is equal to $\lambda_1 + \omega$ from equation (2.2).

[3] See http://www.metoffice.com/research/hadleycentre/models/modeldata.html for the Hadley data. A slight adjustment of the Jones et al. (2005) data was undertaken. The data summarize anomalies relative to the 1961-1990 period. During this period, Jones et al. note that the global average temperature reference value was 14.0°C. The temperature data are transformed by adding this value, calculating the mean for the 1850-1889 period and re-normalizing the data relative to this mean, such that (2.2) is estimated with no constant term.

[4] The historical data are based on the standard IPCC CO_2 concentration history data (Enting et al. 1994), but are used as reported in Joos and Siegenthaler (1999).

Table 1
Regression and ADF test results.

	Regression Results		**ADF Test Results**	
	Jones	HadCM-2	Jones	HadCM-2
G_{t-1}	.6804*	.9399*	-.4642*	-.4288*
	(.06217)	(.02633)	(.07196)	(.08428)
$\frac{\log\left(\frac{m_t}{m_b}\right)}{\log(2)}$	.4031*	.1911*	.9241*	1.721*
	(.08785)	(.06829)	(.1678)	(.3411)
Constant			-.07800*	-.57014*
			(.02172)	(.1249)
Observations	140	100	140	100
Adj R^2	.08361	.9978		
ADF Test Statistic			-6.45	-5.09
			(.000)	(.000)

Standard errors in parentheses (MacKinnon approximate p-values for ADF test statistics).
* indicates statistical significance at the 1% level.

The results of the regressions presented in Table 1 do not allow us to directly determine the expected value of G_{2C}, since this expression is not well-defined over the entire support of the parameter distribution. However, $E[G_{2C}]$ can be estimated by $\frac{\hat{\eta}}{1-\hat{\lambda}}$, although this estimate will be increasingly negatively biased as the standard error of $\hat{\lambda}$ increases, since the value of the transformation is increasing and convex in λ.

While the value of $E[G_{2C}]$ is not well-defined, the Delta method allows for the derivation of the asymptotic covariance matrix for the estimate $G_{2C} = \frac{\hat{\eta}}{1-\hat{\lambda}}$, which is a measure of the state of uncertainty surrounding G_{2C}.[5] Estimates from the Jones et al. (2005) data suggest that:

$$G_{2C} \xrightarrow{a} N[1.283,\ 0.02490],$$

while the HadCM-2 data yield an estimate of

$$G_{2C} \xrightarrow{a} N[3.197,\ 0.1138].$$

[5] For details, see Greene (2003), pp. 108-110.

The empirical exercise above summarizes the ability of information in the statistical record to identify values of climate model parameters and quantifies the nature of uncertainty surrounding these values. To characterize the stepwise updating of these estimates based on the receipt of new information, I use a Bayesian learning framework which takes as given σ_u^2, the variance of the shock in the law of motion for surface temperature. In this section, the full climate model presented in Section 2.1 is used, however uncertainty is assumed to exist only over the values of λ_1 and η, while parameters λ_2 and ω are assumed to be known with certainty. Uncertainty is assumed to exist over these two parameter values in particular since they capture directly the persistence of temperature deviations (λ_1) and the effect of carbon emissions (η) on surface temperature. From a policy perspective, these parameters jointly determine two important measures: the benefit of current emissions control and the expected future temperature as a function of the current climate state.

In order to simplify notation for the learning model, define the temperature change net of the effect of ocean temperature $H_t \equiv G_t - \omega O_{t-1}$ and right-hand side observations:

$$\mathbf{X_t} \equiv \begin{pmatrix} G_{t-1} & \frac{\log\left(\frac{m_{t-1}}{m_b}\right)}{\log(2)} \\ G_{t-2} & \frac{\log\left(\frac{m_{t-2}}{m_b}\right)}{\log(2)} \end{pmatrix}. \tag{2.5}$$

Let the prior distribution be bivariate normal with mean estimates $\mathbf{\Theta} \equiv \begin{pmatrix} \hat{\lambda}_1 & \hat{\eta} \end{pmatrix}'$ and covariance matrix $\mathbf{\Phi}$. Updating based on observations of H and $\mathbf{X}$ in each time period leads to a posterior distribution which is normal with mean $\mathbf{\Theta'}$ and covariance matrix $\mathbf{\Phi'}$, according to updating rules are given as follows:

$$\mathbf{\Theta'} = W\mathbf{\Theta} + (I - W)(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T H, \tag{2.6}$$

where

$$W = \left[\mathbf{\Phi}^{-1} + \left[\sigma_u^2(\mathbf{X}^T\mathbf{X})\right]^{-1}\right]^{-1}\mathbf{\Phi}^{-1} \tag{2.7}$$

and

$$\mathbf{\Phi'} = \left[\mathbf{\Phi}^{-1} + \left[\sigma_u^2(\mathbf{X}^T\mathbf{X})^{-1}\right]^{-1}\right]^{-1}. \tag{2.8}$$

such that the updated estimate of the mean vector in (2.6) is a weighted average of the mean of the prior distribution and the Ordinary Least Squares coefficient vector. [6]

[6] Since two parameter estimates will be updated, there must be at least rows in H and X, such that $X'X$ will be of full rank.

Table 2
Prior Distribution Parameters

	Hadley Prior		**Jones Prior**		**Diffuse Prior**	
	UV	BV	UV	BV	UV	BV
$\hat{\lambda}_1$	0.9112	0.9286	0.9112	0.6711	0.9112	0.9112
SE($\hat{\lambda}_1$)		0.02666		0.06212		0.03252
$\hat{\eta}$	0.3641	0.2838	0.1461	0.5904	0.3416	0.3416
SE($\hat{\eta}$)	0.03842	0.1008	0.01797	0.1289	0.1743	0.1028
Covariance		-0.002642		-0.006633		0
E(G_{2C})	3.192	3.192	1.283	1.283	3	3
Var(E(G_{2C}))	0.1138	0.1138	0.002490	0.002490	2.343	2.343

2.4 Monte Carlo Experiment

In Kelly and Kolstad (1999), an analytic result is presented which defines the expected number of periods it takes for agents to reach a reasonable estimate of the true parameter describing the temperature change induced by accumulated carbon. Below, I use the results of a numerical experiment to further characterize the relationship between carbon emissions, the tightness and accuracy of the prior and the expected number of observations required to reject a false null hypothesis about the values of parameters of the climate system. The experiment results also show the effect of additional uncertainty on the magnitude of potential errors.

The estimator of $E\left[G_{2C}\right] = \frac{\hat{\eta}}{1-\hat{\lambda}_1-\omega}$ is used as the parameter of interest for learning. The experiment uses three prior distributions for which the elements of the mean vectors and covariance matrices are shown in Table 2. The first two priors are the parameter distributions from the regression results reported in Table 1, denoted the Jones and Hadley priors respectively, and the third is a more diffuse prior with accurate mean estimates. For each case, a univariate prior is constructed such that the mean and asymptotic variance of the estimate of G_{2C} are equivalent to those in their corresponding bivariate priors. Subject to the observation of simulated data, posterior distributions are derived from Bayesian updating as described in equations (2.6-2.8). Learning times are reported as the number of observations required, on average, to reject a false null hypothesis about the value of G_{2C} . The experiment proceeds as follows:

Experiment 1 *Monte-Carlo Learning Experiment*

1. *The true parameters of the climate model defined by equations (2.1-2.3) are as given in Table A.1.*[7]
2. *Initial emissions (E_0) are 8.4 GtC and annual emissions (E_t) are determined by growth rate γ_E:*

$$E_t = E_{t-1} * (1 + \gamma_E).$$

3. *From initial values of $m_0 = 770$, $G_0 = 0.31$, and $0_0 = 0.104$, generate 1000 sets of climate data according to equations (2.1-2.3) with $\sigma_u = .11$, given E_t.*
4. *For each of the initial priors given in Table 2, solve for the sequence of posterior distributions for each set of generated data. Updating rules, conditional on simulated data, are given by equations (2.6-2.8).*
5. *Given that the assumed true value for G_{2C} is 3^oC, learning is defined as having occurred when the posterior distribution is such that the null hypotheses H_o:$G_{2C} < 2.9$ and H_o:$G_{2C} > 3.1$ are first simultaneously rejected at the 5% level.*

The results of the experiment characterize the expected number of observations required to extract sufficient information from the statistical record to reject a false null hypothesis about a measure of the severity of climate change. The expected learning time as a function of emissions growth rates for each prior are shown on a logarithmic scale in Figure A.1. Three intuitive results are worthy of note. First, there is a great deal of acceleration in the learning time generated by increases in emissions growth rates. Kelly and Kolstad (1999) discusses this result, noting that emissions control policy is therefore a determinant of expected learning times. Second, learning times increase with the addition of uncertainty over a second parameter. Third, the number of observations required in all cases is such that there is limited sensitivity of learning times to initial prior means, and the tightness of the initial prior has a greater effect on learning times in the single-parameter case.[8] Figure A.2 shows the role of the assumed prior variance in determining learning times. This figure shows the results of a replication of the learning experiment beginning with the diffuse prior means and 2% emissions growth rates for different

[7] These parameter values are standard in the literature. Readers familiar with Nordhaus and Boyer (2000) should note that I assume an annual time interval, not the Nordhaus and Boyer (2000) ten-year interval. Parameter values are thus comparable to Pizer (1999). The assumed true value of η differs slightly from the Pizer (1999) value, as G_{2C} is set to exactly 3^oC for simplicity.

[8] For the case where emissions growth is 2% per year, learning times in the two-parameter case are 206, 210, and 211 years for the diffuse, Jones, and Hadley priors respectively. There is slightly more variation in the single parameter setting, where learning times given the same assumptions are 139, 136 and 161 years respectively.

prior variances.[9] As expected, the learning time decreases non-linearly in the initial variance of the estimate of G_{2C}, *ceteris paribus*.

The results above suggest that there may not be significant differences in the conclusions reached by studying the univariate and bivariate cases, particularly for low emissions growth rates. However, two additional measures of the learning dynamics bring to light important differences; these are the expected learning path and the variance of potential learning paths. The expected learning path informs us as to how much uncertainty would be expected to remain at a point in time in the future for a random draw from the data generating process given by the climate model. Using the asymptotic distribution for parameter estimates $\hat{\lambda}_1$ and $\hat{\eta}$, the expected mean and variance of the estimate of G_{2C} over time are calculated. Figure A.3 shows the evolution of these estimates for the Jones prior.

Two points are well illustrated by this figure. First, the expected estimate of G_{2C} converges more quickly to the true value where the values of two parameters are uncertain. Second, the uncertainty surrounding this estimate is larger and more persistent in the two parameter case, which is a key determinant of the higher learning times reported above. Uncertainty is likely to affect policy choices in that a risk-averse policy-maker would tend to choose lower emissions control, *ceteris paribus*, in the two parameter case since the returns to investing in environmental capital are less certain.

The likelihood that a particular learning path takes us far from the true value is also important, and this is where further differences arise when considering a second unknown parameter. Consider Figure A.4 which shows the bootstrap mean and confidence intervals for the estimate of G_{2C} across samples beginning again from the Jones prior. In the univariate case, the learned value of G_{2C} from any series of data will be within $.265^oC$ of the expected learning path in all time periods 95% of the time. Where two parameter values are uncertain, the learning path can deviate much more significantly from its expected path, lying within 1.36^oC of its expected path 95% of the time. Thus, with two unknown parameters, learning from the statistical record is far more likely to lead to incorrect estimates of G_{2C} than would be the case where the value of a single parameter is uncertain. These results are sensitive to the initial prior variance on one or both parameter estimates. Consider Figure A.5 which shows the same set of results for the diffuse prior. In this case, the bootstrap confidence interval around the estimate of G_{2C} is such that errors of 2.5^oC are possible in the first 30 observations.

[9] Emissions growth rates averaged 2.5% per year over the 1960 to 1998 period, and are expected to average 1.6% over the 1990-2030 period (IEA, 2004). The IS-92a emissions scenario assumed that business-as-usual emissions would grow at 1% per year (IPCC, 2001).

These results show that, where uncertainty exists of the values of two parameters, policy-makers are more likely to face persistent uncertainty and more likely to make significant estimation errors. In what follows, the characterization of learning and uncertainty is imbedded in a model of climate and economy to determine how optimal policy decisions are altered under uncertainty and the potential effects of errors and uncertainty on the evolution of the economy.

3 Integrated Assessment Model

This section embeds the learning characterization with multiple parameter uncertainty outlined above in a modified Nordhaus and Boyer (2000) IAM of the global economy, where a social planner chooses savings and emissions control rates in each period. This extends the model proposed in Kelly and Kolstad (1999) which studies the actions of a social planner facing uncertainty over the effect of current atmospheric carbon stocks on temperature. As compared to the learning experiment in the previous section, results from simulations of this model will allow us to characterize the co-evolution of uncertainty and policy where emissions are an endogenous result of economic activity and the decisions of the planner.

3.1 The Economy

Three sources of exogenous change are assumed to exist in the economy: factor productivity A, the ratio of emissions to output ϕ, and labour supply L, each determined as a function of calendar time T.[10] The generic law of motion for technology $J \in \{A, L, \phi\}$ as a function of time, initial condition J_0, growth rate γ_J, and convergence rates δ_J and $\delta_{2,J}$ is given by:

$$J(T) = J_0 \exp\left(\frac{\gamma_J}{\delta_J}(1 - e^{-\delta_J T})\right). \tag{3.1}$$

Total factor productivity is determined by exogenous factor productivity and two endogenous effects; the choice of emissions control rate $\tau \in [0, 1)$ reduces factor productivity by $(1 - b_1 \tau_t^{b_2})$, and increased surface temperature G reduces aggregate productivity by $(1 + \theta_1 G^{\theta_2})^{-1}$. A single good used for both

[10] Calendar time is an index for the transition of exogenous state variables. It denotes the number of years since the initial values described the state of the economy, and is interpreted as a state variable as in Kelly and Kolstad (1999).

consumption, C, and investment, I, is produced using Cobb-Douglas technology with inputs of capital, K and labour as follows:

$$Y(T, \tau, K, G) = \frac{1 - b_1 \tau^{b_2}}{1 + \theta_1 G^{\theta_2}} A(T) K^{\alpha} L(T)^{1-\alpha} = C + I. \tag{3.2}$$

The capital stock evolves endogenously as a function of chosen investment and depreciation rate $\delta_k \in [0, 1]$ according to:

$$K_{t+1} = (1 - \delta_k) K_t + I_t. \tag{3.3}$$

Emissions E_t are determined by the exogenous ratio of emissions to output and chosen emissions control level, such that

$$E_t = (1 - \tau) \phi(T) Y(T, \tau, K, G). \tag{3.4}$$

Emissions are fed through the climate model described in equations (2.1-2.3) to determine the evolution of state variables G, O, and m.

3.2 Dynamic Optimization

Assume that a social planner maximizes expected welfare through choices of aggregate investment and emissions control rates. Welfare is defined as the expected, discounted stream of population-weighted, per-capita utility, where utility has constant relative risk aversion form. Maintaining the notation for the learning problem used in Section 2.1, denote the state of the economy by $\mathbf{S} = \{K, m, m_{-1}, m_{-2}, G, G_{-1}, O, O_{-1}, T, \mathbf{\Theta}, \mathbf{\Phi}\}$, where j_{-x} and j' denote the x-period lagged and lead values of state variable j respectively. The solution to the planner's recursive problem with learning is characterized using Bellman's equation as:

$$V(S) = \max_{I,\tau} L(T) \left(\frac{\frac{C}{L(T)}}{1-\sigma} \right)^{1-\sigma} + \beta E\Big[V\Big(S'\Big)|\mathbf{\Theta}, \mathbf{\Phi}, I, \tau, \sigma_u^2\Big] \tag{3.5}$$

$$\text{s.t. } Y = \frac{1 - b_1 \tau^{b_2}}{1 + \theta_1 G^{\theta_2}} A(T) K^{\alpha} L(T)^{1-\alpha} \tag{3.6}$$

$$C = Y - I \tag{3.7}$$

$$K' = (1 - \delta_k)K + I \tag{3.8}$$

$$m' = (1 - \tau)\phi(T)Y + (1 - \delta_m)(m - m_b) + m_b \tag{3.9}$$

$$G' = \hat{\lambda}_1 G + \hat{\eta} \left[\frac{\log\left(\frac{m_t}{m_b}\right)}{\log(2)} \right] + \omega O + u,\ u \sim N(0, \sigma_u^2) \tag{3.10}$$

$$O' = \lambda_2 O + (1 - \lambda_2) G \tag{3.11}$$

$$\mathbf{\Theta}' = W\mathbf{\Theta} + (I - W)(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T H \tag{3.12}$$

$$\mathbf{\Phi}^* = \Big[\mathbf{\Phi}^{-1} + \Big[\sigma_u^2 (\mathbf{X}^T\mathbf{X})^{-1}\Big]^{-1}\Big]^{-1} \tag{3.13}$$

$$T' = T + 1. \tag{3.14}$$

The parameters of the model are $\{A_0,\ L_0,\ \phi_0,\ b_1,\ b_2,\ \theta_1,\ \theta_2,\ \lambda_2,\ \delta_k,\ \delta_m,\ \delta_a,\ \delta_l,\ \delta_\phi,\ \gamma_a,\ \gamma_l,\gamma_\phi,\ m_b,\ \lambda_1,\ \eta,\ \omega,\ \sigma_u^2\}$. Note that the true λ_1 and η values are used only to simulate the path of the learning economy, while the planner's decisions are based on the elements of $\mathbf{\Theta}$; parameter estimates $\hat{\lambda}_1$ and $\hat{\eta}$.

4 Computation

Solving the recursive problem described in equations (3.5-3.14) presents a challenge of dimensionality. In order to solve the model, small modifications are made to the technique described in Kelly and Kolstad (1999, 2001), which approximates the solution using an iterative algorithm combined with the use of a neural network to approximation of the value function over a finite set of grid points.

The neural network approximation is defined as follows.[11] Define by L the number of nodes in the hidden layer, and let n represent the number of state variables in the model, such that the state space is $\mathbb{R}^n$. Denote by $\mathbf{x} \in \mathbb{R}^{n+1}$ a set of real-valued signals to the network, with the first element ($x_0 = 1$) being a bias signal, and the remaining elements being the state vector for a particular element of the state space. Let χ_1 be a $(n+1) \times L$ matrix of inner weights, and let $\mathbf{z}(\mathbf{x}, \chi_\mathbf{1})$ be a $(L+1) \times 1$ vector, with the first element ($z_0 = 1$) being a bias signal. The additional elements ($z_1..z_L$) are the output

[11] For a detailed discussion of neural networks, the interested reader is referred to Hassoun (1995).

Table 3
State variable ranges used in solving the model.

State Variable	**Definition**	**Units**	**Grid** Min	Max	1995 Value [a]
K	Capital Stock	$10^{12}\$US_{1997}$	10	1200	55.75
m	Atmospheric Carbon	GtC	590	1800	770
G	Global Surface Temp.	^oC	-.75	12	.31
O	Ocean Temp.	^oC	0	10	.104
G_{-1}	Temp. Lag 1	^oC	0	12	.24
$\hat{\lambda}_1$	Mean Estimate of λ_1	-	.5	.97	.9112
$\hat{\eta}$	Mean Estimate of η	^oC	0	.7	.34163
$\mathbf{\Phi}_{11}$	Variance Estimate $\hat{\lambda}_1$	-	0	.02	1^{-10}
$\mathbf{\Phi}_{22}$	Variance Estimate $\hat{\eta}$	$(^oC)^2$	0	.2	1^{-10}
$\frac{\mathbf{\Phi}_{12}}{\sqrt{\mathbf{\Phi}_{11}}\sqrt{\mathbf{\Phi}_{22}}}$	Correlation of Estimates	-	-1	1	0
G_{2C}	Long run G for $m = 2 \times m_b$	^oC	0	8	3
T	Technology Index	years	0	600	0

[a] From the calibration of the model.

values from the hidden layer of the network, $\mathbf{z_l} = \tanh(\chi'_{1l}\mathbf{x})\,\forall l = 1..L$, where χ_{1l} is a column of χ_1. The $L+1$ elements of $\mathbf{z}(\mathbf{x}, \chi_\mathbf{1})$ are then aggregated using outer weights χ_2, a $(L+1) \times 1$ vector. The approximation is expressed as:

$$\Upsilon(\mathbf{x}) = \chi'_2\left(z(\mathbf{x}, \chi_1)\right). \tag{4.1}$$

Using $\Upsilon(\mathbf{x})$ as an approximation to the value function defined over $\mathbb{R}^{n+1}$, the following iterative algorithm provides an approximate solution to the dynamic programming problem.

Algorithm 1 *Algorithm Preliminaries: Choose a convergence criterion ϵ, number of network nodes L, and starting values for the weights χ_1 and χ_2 in $\Upsilon(\mathbf{x})$. Draw N grid points from a low discrepancy sequence and transform these points such that the bounds of the state space are those in Table 3.* [12]

Step 1: For each point on the grid $\{s_i\}_{i=1}^N$, define $\mathbf{x_i} = (1\ s_{i,1}..s_{i,n})$, and solve the maximization problem given in (3.5-3.14). Use analytic derivatives of the approximating function to solve the first order conditions for optimal I_i^* and

[12] A Halton sequence is used to draw a set of grid points which are uniformly distributed within the state space. For details on low-discrepancy sequences, see Judd (1998). For the results reported below, $\epsilon = 10^{-4}$, L=24, and N=100 000.

τ_i^*.
Step 2: Denoting by $\mathbf{x}'$ *the value of* $\mathbf{x}$ *corresponding to future states, update the value at each point on the grid as:*

$$V^{j+1}(s_i) = U(s_i|I_i^*,\, \tau_i^*) + \beta E\Upsilon(\mathbf{x_i}'|\mathbf{x_i},\, I_i^*,\, \tau_i^*) \tag{4.2}$$

Step 3: Use updated values $\{V^{j+1}(s_i)\}_{i=1}^N$ *to solve for new weights* χ_1 *and* χ_2 *that minimize* $\|V^{j+1}(s) - \Upsilon(\mathbf{x})\|$.

Step 4: Return to Step 1 unless $\|\frac{V^{j+1}(s)-V^j(s)}{V^j(s)}\| < \epsilon$.

Four changes are made to the algorithm detailed in Kelly and Kolstad (2001). First, grid points are drawn from a low-discrepancy sequence. A low-discrepancy sequence is an analog to a multi-dimensional, uniform random number generator. The low-discrepancy sequence minimizes the maximum Euclidean distance between two grid points to generate the most even coverage of points within a hypercube, which makes it an ideal tool for this problem. Three modifications are made to the low-discrepancy draw. First, values for the capital stock and time period are re-scaled such that there is more density near the lower bounds. Second, values of $\hat{\eta}$ are re-scaled such that the value G_{2C} does not exceed 8^oC. Finally, draws of G are used to assign values to state variable G_{-1} such that lagged temperature cannot be outside the possible values implied by the ranges of η and λ_1 on the grid. These assumptions allow grid points to be consistent with the model and concentrated in areas where the combinations of $\hat{\eta}$ and $\hat{\lambda}_1$ are consistent with the potential magnitude of global climate change, or where substantial curvature is likely to exist in the value function.

The second modification to the Kelly and Kolstad (2001) algorithm is the use of Monte-Carlo integration rather than quadrature-based methods. Within each pass of the value function iteration, expected future values are evaluated using 100 point draws from the distribution of possible future states conditional on the prior distribution, the variance of temperature shocks, and the choice variables for investment and emissions control.

In order to render the problem solvable, and still capture the nature of the policy problem at hand, it is necessary to make an assumption about the planner's expectations. Draws from the distribution of possible future states are censored such that they lie within the bounds of the state grid. As such, the planner does not admit the possibility of non-stationary temperature processes, or combinations of η and λ_1 which imply $G_{2C} > 8$. This will bias downward the planner's expectations over future temperatures, and bias upward their evaluation of the effect of emissions reduction on future temperature changes. Additionally, the time period is unbounded, however its transition is censored such that its value does not exceed 600 periods. For state points close to the boundaries of the grid, these assumptions will introduce some bias to the solution.

The third modification to the Kelly and Kolstad (2001) algorithm is that to solve the non-linear, minimization problem in *Step 3*, a sample of 25000 observations is randomly drawn, with replacement, at each iteration. Thus, computation time is limited by not solving the approximation over all grid points at each iteration, but all grid points have equal expected leverage on the parameters of the neural network approximation.

Finally, in order to restrict the number of state variables and render the problem tractable, an assumption is made on the learning dynamics. The planner is assumed not to keep track of the lagged values of either ocean temperature or atmospheric carbon. Instead, when she enters a new period and observes the evolution of the state, she is assumed to calculate growth rates for ocean temperature (γ_O) and atmospheric carbon forcing (γ_f) based on the previous state of the economy, and use those growth rates to impute lagged values for forcing and ocean temperature, such that some elements of the X and H matrices are imputed as:

$$\mathbf{X} = \begin{pmatrix} G & \log \frac{m}{m_b} \\ G_{-1} & \frac{\log \frac{m}{m_b}}{(1+\gamma_f)} \end{pmatrix}, \quad H = \begin{pmatrix} G' - \omega O \\ G - \omega \frac{O_t}{(1+\gamma_O)} \end{pmatrix}.$$

The imputation above assumes that the second derivatives of O and m are zero over two periods to allow the number of state variables in the planner's problem to be reduced while maintaining full rank for the matrix of observations used for learning. Lagged temperature is not treated in the same way, since the second derivative of temperature at any point is a function of previous draws of the random shock, and as such not necessarily close to zero.

5 Calibration

To calibrate the economy, I begin by setting most parameter values to values from Kelly and Kolstad (1999) and Pizer (1999). The exceptions to this are as follows. The variance of temperature residuals is taken from the estimation of the regression model using Jones et al. (2005) temperature data reported in Table 1. Starting values for each of the technology levels are set to 1995 values (ie. T=0 implies the year 1995). The initial labour supply and its growth parameters are set such that its transition matches the median population growth scenario from United Nations (2004) for years 1995-2050. The law of motion for the ratio of emissions to output is calibrated to match the same ratio in IEA (2004) for years 1995-2030. The initial capital stock is the Pizer (1999) value of $\$US_{1997}55.75 \times 10^{12}$.[13] Finally, the trend parameters for ex-

[13] The currency conversion used is the ratio between the Pizer (1999) value for 1995 output, $\$US_{1989}24 \times 10^{12}$ and the IEA value for the same period of $\$US_{1997}34.234 \times$

Table 4
Policy choice correlation with state variables, rounded to two decimal places.

	K	m	G	G_{-1}	O	$\hat{\lambda}_1$	$\hat{\eta}$	$\mathbf{\Phi}_{11}$	$\mathbf{\Phi}_{22}$	$\frac{\mathbf{\Phi}_{12}}{\sqrt{\mathbf{\Phi}_{11}}\sqrt{\mathbf{\Phi}_{22}}}$	T
I	-0.23	0.01	-0.27	-0.27	-0.00	0.03	-0.01	-0.00	-0.00	-0.01	0.79
τ	0.68	0.11	0.24	0.17	0.02	0.41	-0.11	-0.03	-0.07	0.05	0.14

ogenous factor productivity are chosen using a simulation of the economy. I use an approximate certainty benchmark where the planner's prior has correct mean estimates, the diagonal elements of the covariance matrix are set to 10^{-10}, and the correlation of estimates is set to zero. This approximation is necessary since the model does not actually nest perfect certainty, as the covariance would be undefined if either variance term were set to exactly zero. [14] The parameters for factor productivity are chosen such that the initial conditions of the model and short-run transitions under the benchmark assumptions match International Energy Agency (2004) output data and projections for 1995-2030. Starting values used for calibration are shown in Table 3, and final parameter values are shown in Table A.1. The ability of the calibrated model to match medium-term trends in population, aggregate production, and emissions is shown in Figure A.6.

From the solution to the social planner's problem, the influence of state variables on policy choices can be established conditional on the calibration. The unconditional correlations between the choice and state variables of the planner's problem are shown in Table 4. The results are largely intuitive; the effects of risk-aversion, decreasing marginal returns to capital, temperature are as expected. It is also shown that the planner's beliefs on the both the persistence of temperature deviations and the severity of climate change are important for the choice of τ. The counter-intuitive correlations for $\hat{\eta}$ are a result of the negative correlation between $\hat{\lambda}_1$ and $\hat{\eta}$ implied by the G_{2C} constraint on the state space. The correlations between $\hat{\eta}$ and I and τ for larger values of $\hat{\lambda}_1$ are positive.

6 Simulations

Having solved the social planner's problem for the calibrated model, simulations of the transition path characterize the effects of uncertainty and learning on economy. In particular, I show the effect of endogenous emissions control on

10^{12}.

[14] I maintain the same assumption for implementing the univariate priors below. The variance of the estimate of λ_1 is set to 10^{-10}, and the correlation of estimates is set to the value from the bivariate prior.

learning dynamics, and the effect of uncertainty on emissions control choices. The simulations are draws from the economy characterized in Section 3. For each simulation, the true parameter values and initial state of the model are fixed. In each time period, the optimal decisions of the social planner given the economic and climate states and a draw from the temperature shock distribution determine the transition of the economy. The simulations are repeated 1000 times for each set of initial conditions to yield the average transition paths and confidence intervals reported below. Three specific results are reported. First, the expected learning paths for the planner, and the confidence intervals around these paths are presented. Second, the sensitivity of the learning path dynamics to the assumed true value of G_{2C} is tested. Finally, the effects of uncertainty and learning dynamics on emissions control policy are shown. Where applicable, the results are reported relative to the benchmark model used for calibration.

The initial state variables are the 2005 values from the calibration run of the economy, such that all learning and policy choices take place in the future. Simulations use the univariate and bivariate versions of the Jones, Hadley, and diffuse priors described in Table 2, although recall that I approximate the univariate priors by setting $\hat{\lambda}_1$ to its true value and setting its variance to 10^{-10}. The results under the diffuse, univariate prior provide a close reference to the IAM results in Kelly and Kolstad (1999). [15]

As shown in Section 2.4, the growth rate of emissions is an important determinant of learning dynamics. For a given emissions profile, the learning dynamics in the IAM and the experiment will be almost identical, with slight differences coming from the assumption of imputation of historical data by the planner and the bounds on the learning state. However, emissions growth rates are endogenous, stochastic, and decreasing over time in the IAM, whereas the experiment uses constant, pre-determined growth rates. Uncertainty should have an effect on learning times since the risk aversion of the planner should lead to decreased emissions control, and thus decreased learning times the greater the level of uncertainty. The link between climate change and the rate of return to future physical capital could also lead to growth effects of uncertainty, although the near-zero correlation between investment and expectations of future climate change reported in Table 4 suggests this effect is not likely to be important. Furthermore, since errors in estimates of the severity of climate change are more likely to arise where two parameters of the system are unknown and emissions control choices will depend on these estimates, uncertainty over multiple parameter values should increase the variance in policy

[15] In particular, Kelly and Kolstad (1999) uses an initial prior with the correct estimate of G_{2C} (2.5^oC in the paper) and a variance of this estimate of 7.158. The diffuse priors I use also provide a correct initial estimate but with a slightly tighter variance.

choices across simulations.

Figure A.7 shows the expected learning path for the planner, as well as the bootstrap confidence intervals for these learning paths for the three priors. Clearly the dynamics discussed earlier in the experiment are replicated in the IAM. The univariate priors lead to slower transitions toward a correct estimate of G_{2C} in expectation, however the confidence intervals describing possible learning curves are very tight around the expected path. The bivariate priors lead to better estimates in expectation, with larger potential errors.

The results above depend on the calibration of G_{2C} to 3^oC. Figure A.8 shows the expected learning paths and confidence intervals for the bivariate, diffuse prior where the assumed true value for η is varied such that G_{2C} is either 4.5^oC or 1.5^oC. Clearly, the resolution of uncertainty varies both in expectation and in the magnitude of potential errors depending on the sign of the error in initial beliefs. The expected learning path reaches the true value more quickly where the true G_{2C} is 4.5^oC, although there is a slight over-shooting. In the case of an initial over-estimate of G_{2C} , there is a slower adjustment of the estimate toward the true value. Where initial beliefs underestimate (overestimate) the severity of climate change, the likelihood of large errors in the estimate of G_{2C} increases (decreases) significantly.

Uncertainty and estimate errors each affect the planner's choice of emissions control. Figure A.9 shows the mean choice of emissions control policy for each of the bivariate priors and the implied under-investment in emissions control along each transition path relative to the calibration benchmark. The initial underestimate of G_{2C} implied by the Jones prior leads to sustained under-investment over the entire simulation period. While the estimate of G_{2C} is improving over time, benchmark emissions control and GDP are increasing as well. Since the costs of regulation are measured in terms of factor productivity, even though the gap between the benchmark policy and the policy choices under the Jones prior is decreasing over time, the amount of under-investment is increasing. The Hadley prior adjusts only very slightly over the sample period, maintaining a slight over-estimation of G_{2C} and therefore a slight over-investment in emissions control.

The role of uncertainty, *ceteris paribus*, is demonstrated by the transitions from the diffuse priors. The planner's parameter value estimates are correct in expectation, however expected investment in emissions control is slightly below the benchmark. While not shown graphically, simulations with the diffuse prior also confirm that spreading uncertainty over the values of two parameters has no effect in and of itself. Differences in chosen policy are insignificant between the univariate and bivariate implementations of the diffuse prior, which share identical and correct estimates of underlying parameter values in expectation. However, as expected, there is a slight effect of spreading uncertainty over two

parameter values on the variance of chosen policy. In each of the simulations, the variance in policy choices across samples is less than 10^{-5}, but three to ten times higher in the bivariate implementations of each prior than for the corresponding univariate prior.

The motivation for this analysis is that the interplay between natural and anthropogenic sources of climate change is important for choosing optimal policy. For a given estimate of G_{2C} , the planner's choice of emissions reduction will be a function of the relative importance of estimates $\hat{\eta}$ and $\hat{\lambda}_1$. This effect is clearly shown in Figure A.10, which plots the expected learning curve for both parameter estimates along with policy choices under the Jones priors. Despite the fact that the mean estimate of G_{2C} is higher throughout the transition for the bivariate prior, chosen regulation is lower throughout. This result is due to an under-estimate of the persistence of climate changes in the bivariate prior, which has initial $\hat{\lambda}_1 = .6711$, as compared to the correct estimate $\hat{\lambda}_1 = .9112$ in the univariate case. While the estimate in the bivariate prior adjusts over time, it remains below the true value. Thus, while the planner estimates a higher marginal effect of emissions on next period's temperature under the bivariate prior, the lack of persistence is such that the marginal welfare generated through emissions control will be smaller. This result is of particular importance since it demonstrates clearly the leverage that λ_1 has on the marginal benefit of emissions control, and, as a consequence, the effects of uncertainty over the value of this parameter on the likelihood of errors in policy.

7 Conclusion

This paper uses both a reduced-form, numerical experiment and a dynamic, optimal policy model to explore the effects of learning and uncertainty on the ability to set effective climate change mitigation policies. This paper deals specifically with that fact that it is difficult to determine the benefits of emissions control when the relative importance of natural trends and anthropogenic influences on temperature changes are unknown.

This paper extends earlier results from Kelly and Kolstad (1999) to show that where uncertainty exists over two potential causes of observed climate changes, the time to learn the true parameter values of the climate model may be on the order of hundreds if not thousands of years. Perhaps more importantly, it is shown that the probability that a particular learning path yields very poor estimates of the severity of climate change is greatly affected by the nature of initial uncertainty, and that uncertainty can be expected to be much more persistent where more parameter values are uncertain. In part, the results suggest that some trade-off must be made between investments

in regulation under uncertainty and investments in accelerating the arrival of new information. Further, the results emphasize the fact that temperature data represent only a single draw from a complicated system about which we have limited knowledge. In such an environment, allowing this single set of observations to over-influence our policy choices may lead to significant errors.

The results presented here are limited by computational complexity. Uncertainty clearly exists over values which are as important for the determination of effective climate policies as those explored in this paper: the extent of damages, the cost of reducing emissions, and future technological progress are but three important examples. Furthermore, even greater uncertainty may exist over the regional distribution of the effects of climate change and potential thresholds and irreversibility in the climate system, which are also not treated in this paper.

References

[1] I.E. Enting, T.M.L. Wigley, and M. Heimann, *Future Emissions and Concentrations of Carbon Dioxide: Key Ocean/Atmosphere/Land Analyses*, CSIRO Division of Atmospheric Research Technical Paper No. 31., 1994.

[2] William H. Greene, *Econometric Analysis*, Fifth ed., Prentice Hall, Upper Saddle River, N.J., USA, 2003.

[3] Mohamad H. Hassoun, *Fundamentals of Artificial Neural Networks*, The MIT Press, Cambridge, Massachusetts, USA, 1995.

[4] International Energy Agency, *World Energy Outlook: 2004*, IEA Publications, Paris, 2004.

[5] T. C. Johns, R. E. Carnell, J. F. Crossley, J. M. Gregory, J. F. B. Mitchell, C. A. Senior, S. F. B. Tett, and R. A. Wood, *The Second Hadley Centre coupled ocean-atmosphere GCM: Model description, spinup and validation*, Climate Dynamics **13** (1997), 103–134.

[6] P.D. Jones, D.E. Parker, T.J. Osborn, and K.R. Briffa, *Global and hemispheric temperature anomalies – land and marine instrument records.*, In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A., 2005.

[7] F. Joos and U. Siegenthaler, *IPCC 1992 scenarios: CO2 concentration (ppm)*, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN, USA, 1999.

[8] Kenneth L. Judd, *Numerical Methods in Economics*, Massachusetts Institute of Technology, Cambridge, Mass., USA, 1998.

[9] Larry Karp and Jiangfeng Zhang, *Regulation with anticipated learning about environmental damages*, Revised, 2004, 2002.

[10] David L. Kelly and Charles D. Kolstad, *Bayesian learning, growth and pollution*, Journal of Economic Dynamics and Control **23** (1999), 491–518.

[11] ———, *Solving growth models with an environmental sector*, Journal of Computational Economics **18** (2001), 217–235.

[12] David L. Kelly, Charles D. Kolstad, and Glenn Mitchell, *Adjustment costs from environmental change induced by incomplete information and learning*, Working Paper, 1999.

[13] Charles D. Kolstad, *International Environmental Economics: Theories, Models and Applications to Climate Change, Internation Trade and Acidification*, ch. The Timing of CO_2 Control in the Face of Uncertainty and Learning, pp. 75–96, Elsevier Science, 1994.

[14] ———, *Learning and stock effects in environmental regulation: The case of greenhouse gas emissions*, Journal of Environmental Economics and Management **18** (1996), 1–18.

[15] ———, *The economics of global warming* , The International Library of Critical Writings in Economics, ch. The Timing of CO_2 Control in the Face of Uncertainty and Learning, pp. 523–544, Elgar Reference Collection, Cheltenham, UK, 1997.

[16] Alan S. Manne, Robert Mendelsohn, and Richard Richels, *MERGE: A model for evaluating regional and global effects of GHG reduction policies*, Energy Policy **23** (1995), no. 1, 17–34.

[17] Alan S. Manne and Richard Richels, *Buying Greenhouse Insurance: The Economic Costs of CO_2 Emissions Limits*, The MIT Press, Cambridge, Mass., 1992.

[18] William D. Nordhaus, *Managing the Global Commons: The Economics of Climate Change*, MIT Press, Cambridge, Mass., 1994.

[19] William D. Nordhaus and Joseph Boyer, *Warming the World*, MIT Press, Cambridge, Mass., 2000.

[20] William A. Pizer, *The optimal choice of climate change policy in the presence of uncertainty*, Resource and Energy Economics **21** (1999), 255–87.

[21] Michael E. Schlesinger, Natalia G. Andronova, Charles D. Kolstad, and David L. Kelly, *On the Use of Autoregression Models to Estimate Climate Sensitivity*, undated.

[22] Alistair Ulph and David Ulph, *Global warming, irreversibility and learning*, The Economic Journal **107** (1997), 636–650.

[23] United Nations Population Division, *World Population Prospects Population Database*, United Nations, New York, USA, 2004.

Table A.1
Calibrated Values

Parameter	Description	Calibrated Value
Inter-temporal Utility Function		
σ	Coefficient of risk aversion	1.2213
β	Discount factor	.9524
Production, Technology, and Labour Supply		
α	Production share of capital	.25
δ_k	Capital depreciation rate	.1
A_0	Initial factor productivity	.01917
γ_a	Initial growth rate of factor productivity	.0195
δ_a	Rate of decline of γ_a	.0118
L_0	Initial labour supply	5674
γ_L	Initial growth rate of labour supply	0.015
δ_L	Rate of decline of γ_L	.024
ϕ_0	Initial emissions: output ratio	0.166992
γ_ϕ	Initial growth rate of emissions: output ratio	-0.013
δ_ϕ	Rate of decline of γ_ϕ	.00161164
Climate		
m_b	Preindustrial carbon stock (GtC)	590
$(1-\delta_m)$	Atmospheric retention of carbon	.99167
λ_1	AR component of temperature change	.9112
λ_2	AR component of ocean temperature change	.98
η	Forcing parameter (^oC)	.34163
ω	Transfer between ocean and surface temperature	.009866
G_{2C}	Long-run temperature change for $m = 2m_b$ (^oC)	3
σ_u	Standard deviation of temperature residuals	.11
Emissions Control Technology		
b_1	Linear control costs	.0686
b_2	Exponential control costs	2.887
Productivity Loss from Climate Change		
θ_1	Linear component of damages	.00071
θ_2	Exponent in damage function	.00242

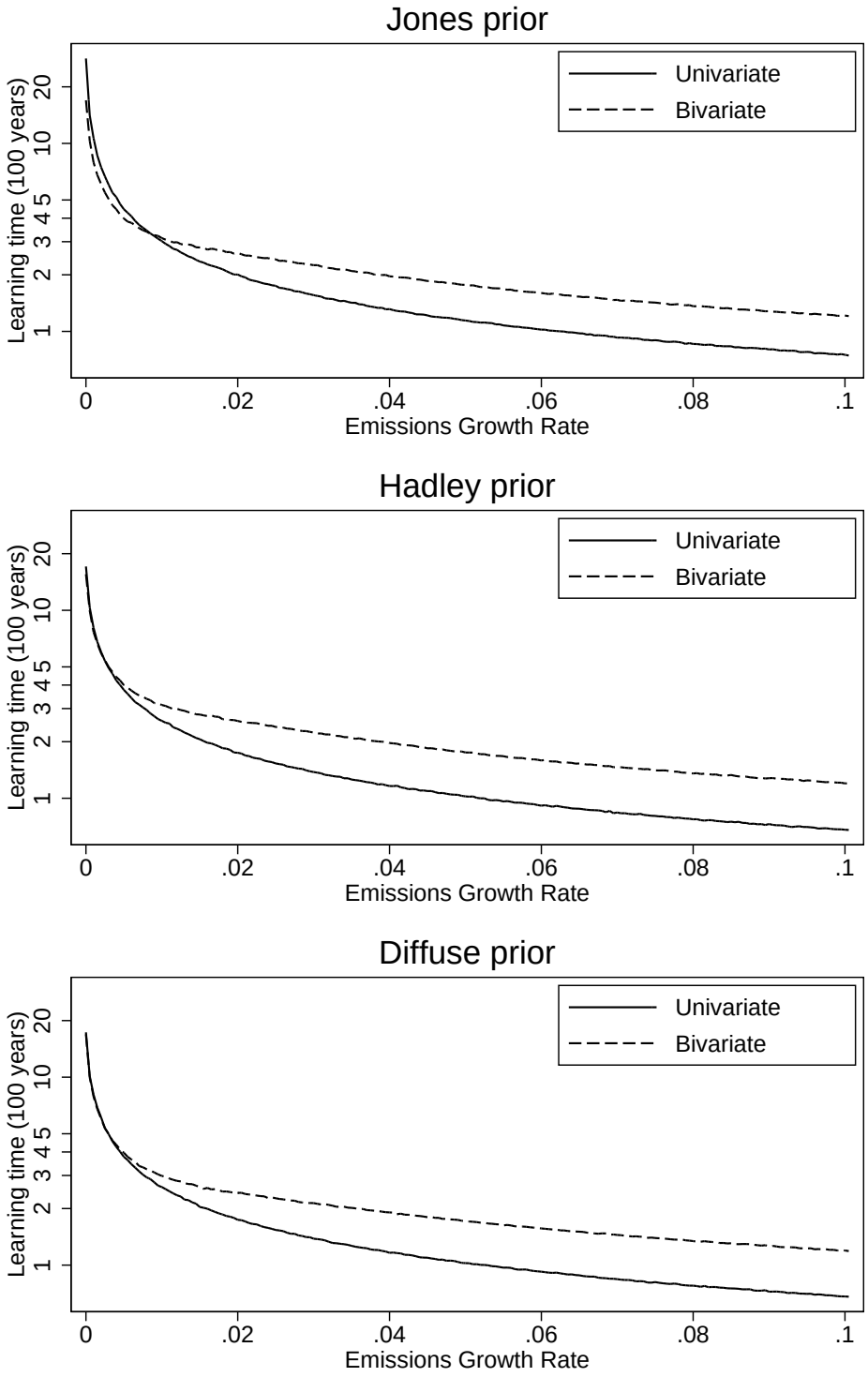

Fig. A.1. Learning Time (log scale) vs. Growth Rate of Emissions. The curves show the number of time periods required before we are able to reject the hypotheses H_o:$G_{2C} < 2.9$ and H_o:$G_{2C} > 3.1$ at the 5% level, beginning from each of the priors. Exogenous growth rates of emissions (γ_E) are shown on the horizontal axis.

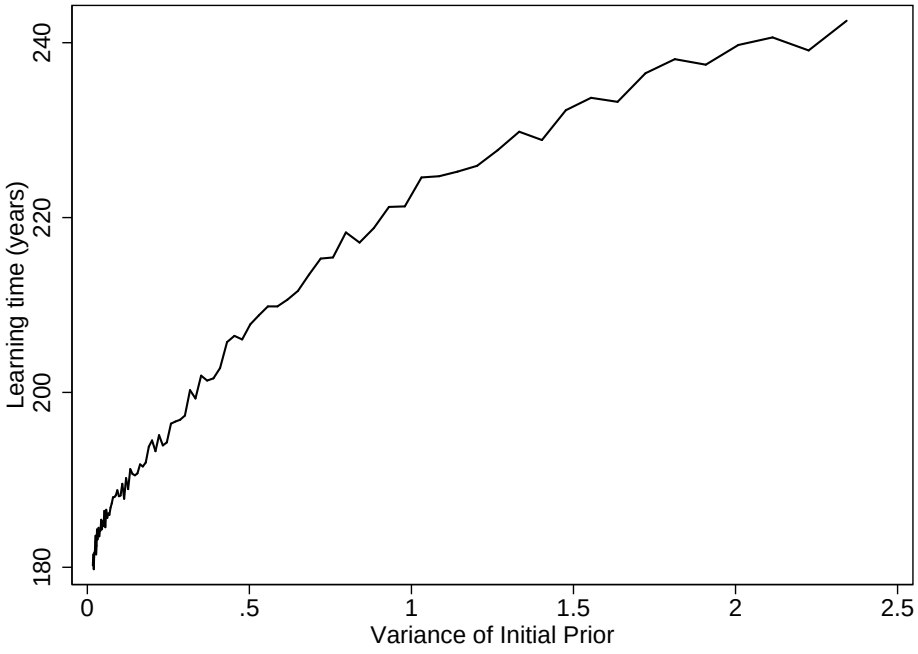

Fig. A.2. Learning times vs. initial variance in G_{2C} estimate. In this Figure, the learning times are shown for the bivariate diffuse prior with progressively tighter variance around the estimate of G_{2C} .

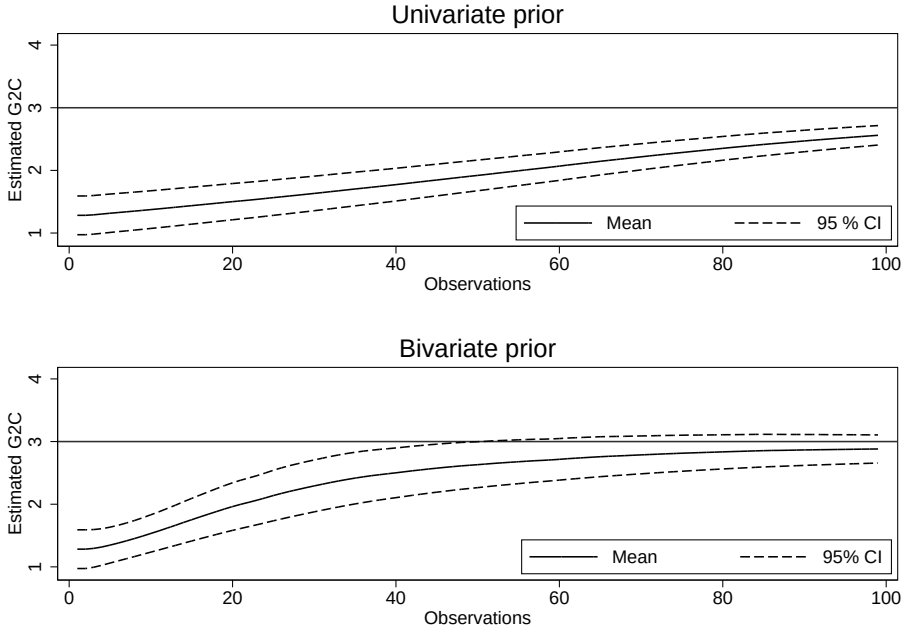

Fig. A.3. G_{2C} estimate and 95% confidence intervals implied by expected parameter distribution over time, Jones prior.

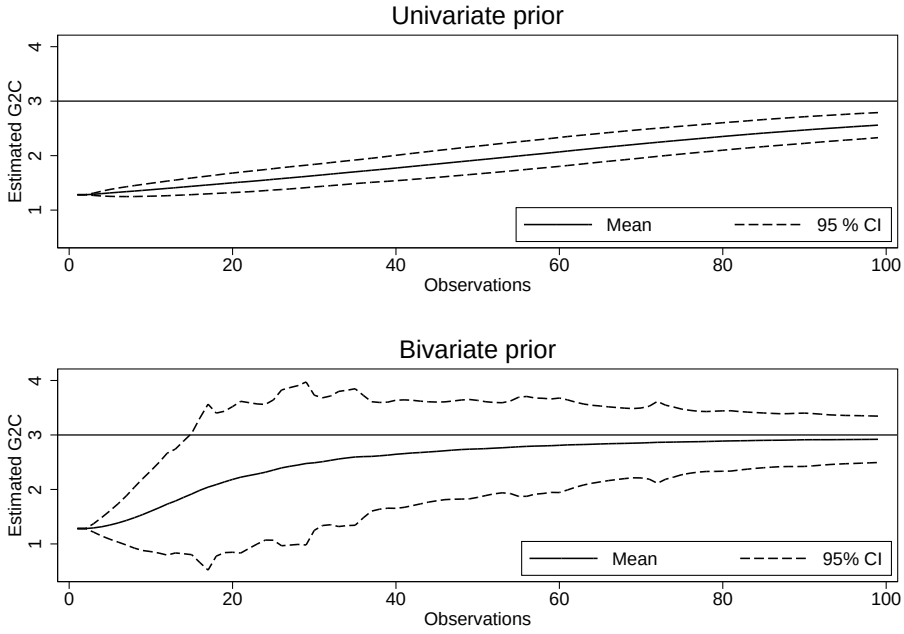

Fig. A.4. Bootstrap mean and confidence intervals for estimates of G_{2C} over time, Jones prior.

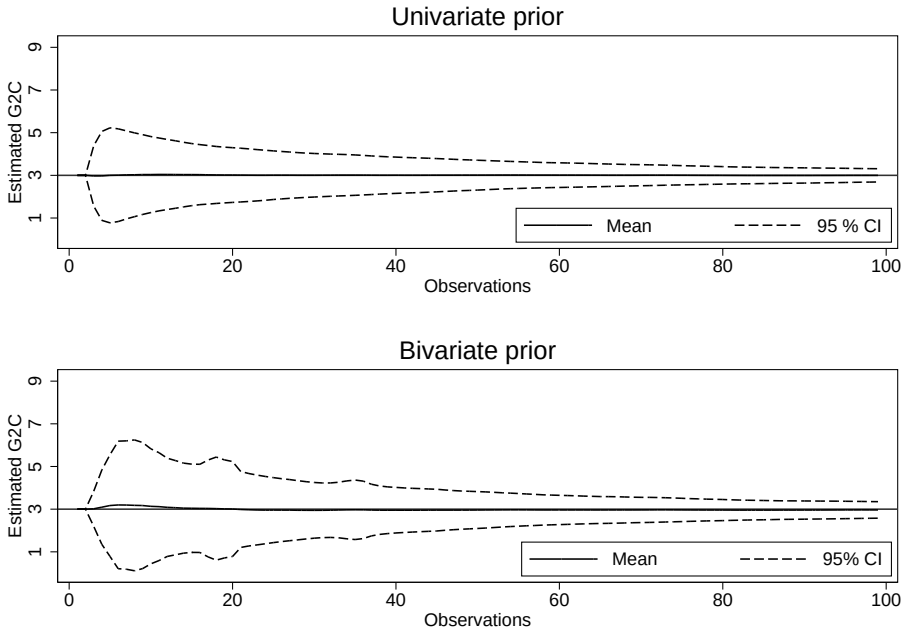

Fig. A.5. Bootstrap mean and confidence intervals for estimates of G_{2C} over time, diffuse prior.

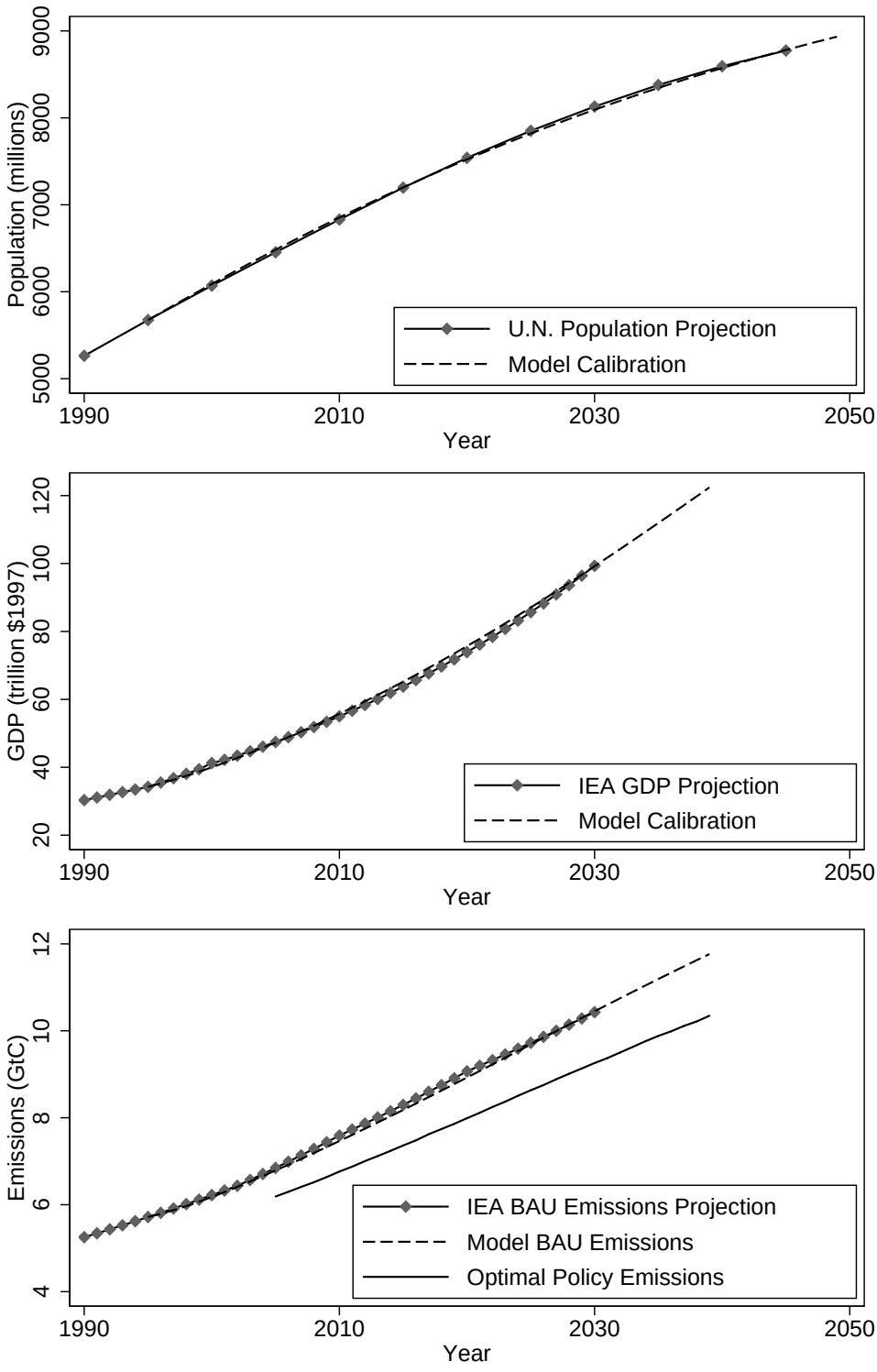

Fig. A.6. Model calibration for labour supply, GDP, and emissions.

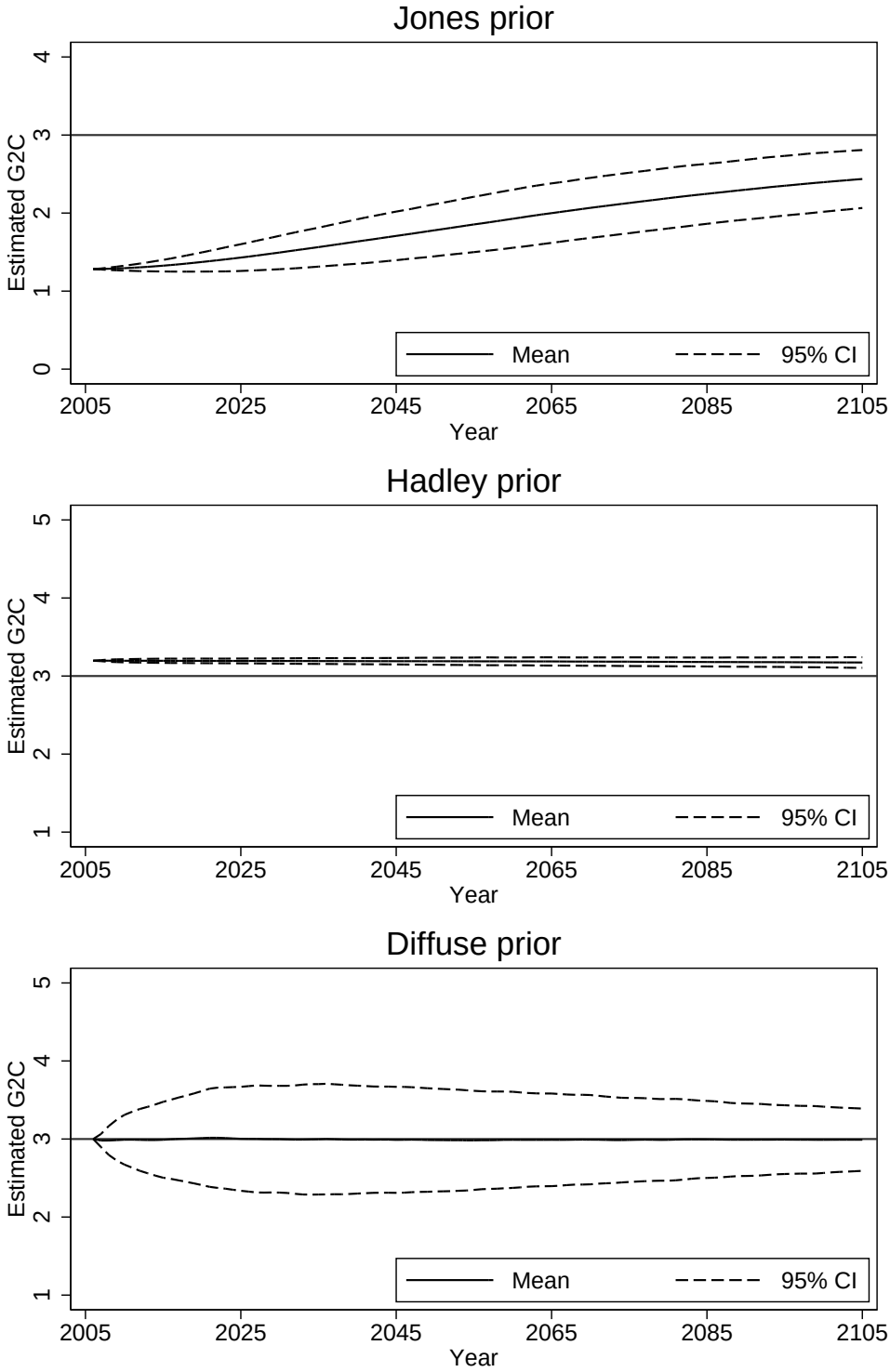

Fig. A.7. Planner's learning curve: Bootstrap mean and confidence intervals for planner's G_{2C} estimate over time

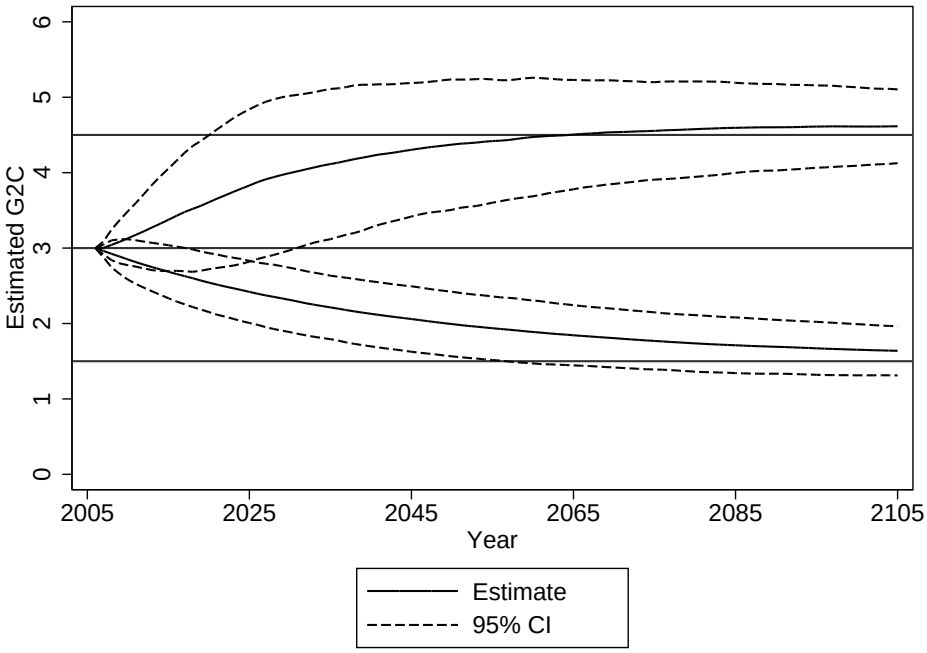

Fig. A.8. Bootstrap mean and confidence intervals for planner's G_{2C} estimate over time. In this Figure, η is re-calibrated such that G_{2C} is either 4.5^{o}C or 1.5^{o}C. Initial priors are the bivariate, diffuse prior.

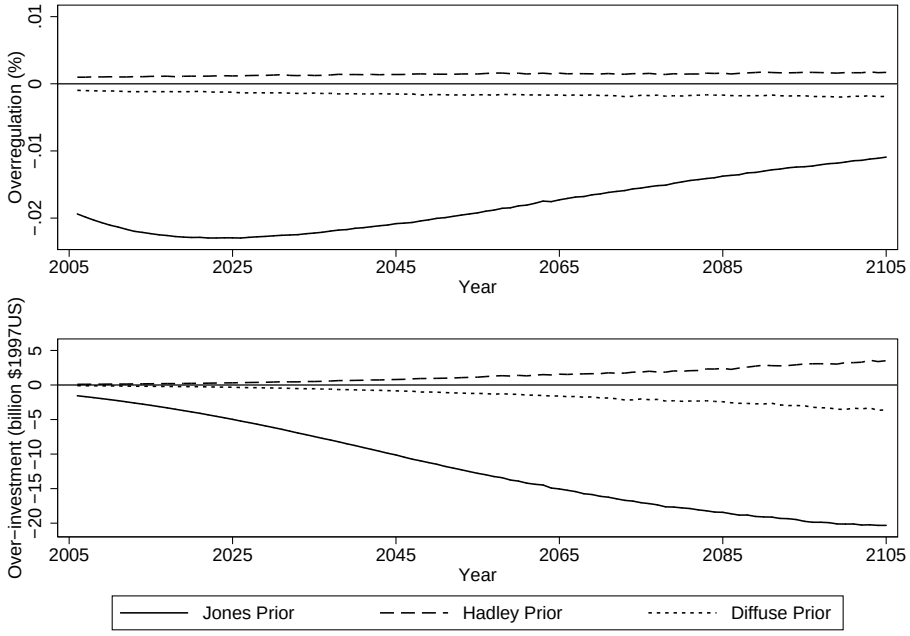

Fig. A.9. Difference between mean policy choice under Jones, Hadley, and diffuse priors and policy choices under the certainty benchmark. The units for difference in emissions control are the absolute reduction in the percentage reduction over uncontrolled emissions.

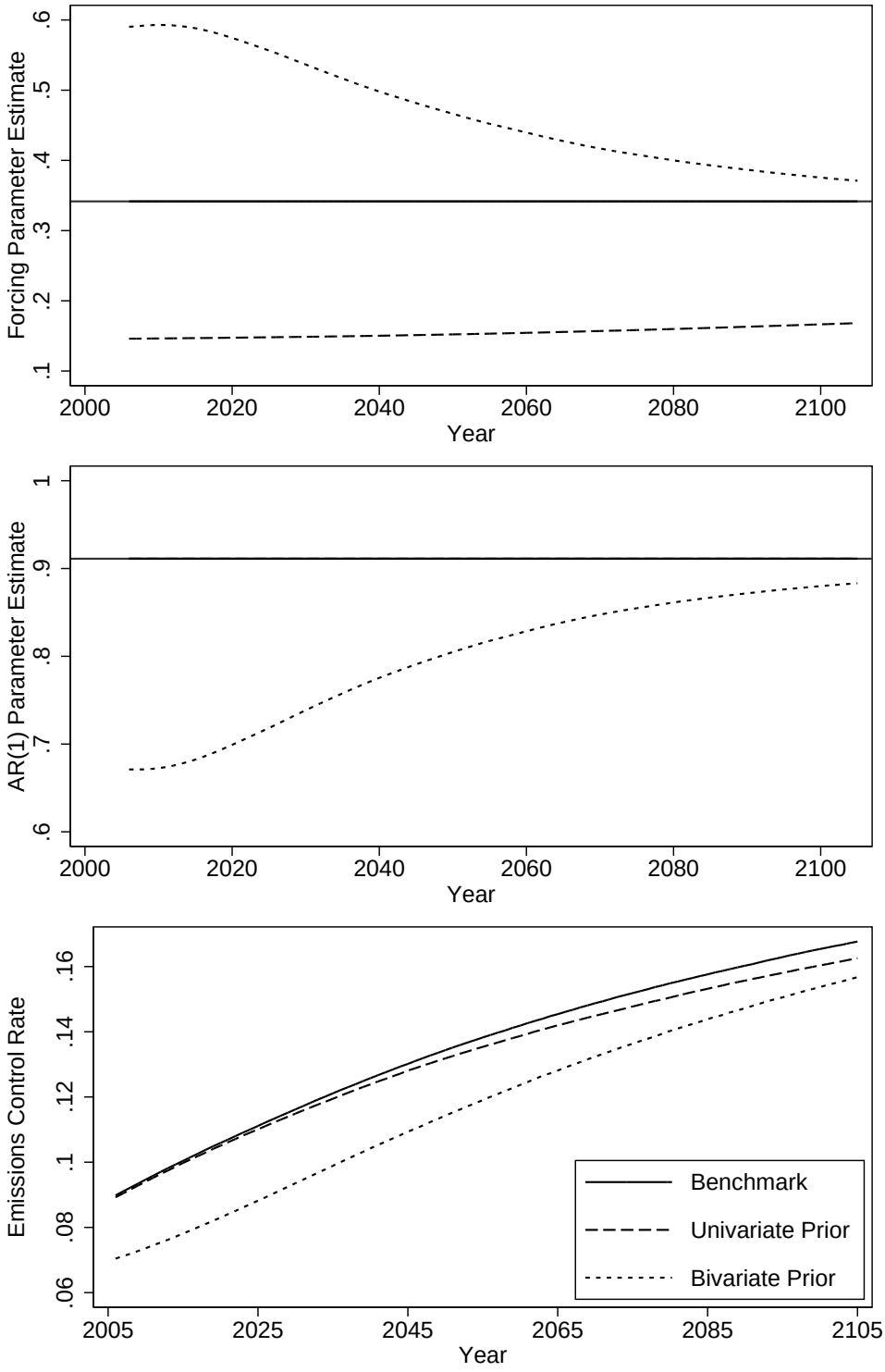

Fig. A.10. Policy choices and learning, Jones prior. This Figure shows the average choice of emissions control along with values of parameter estimates learned over the transition path beginning from the Jones priors compared to the benchmark.