SNF Report No. 16/11

A stitch in time saves nine
The costs of postponing action in climate policy

by

Stein Ivar Steinshamn
Sturla F. Kvamsdal
Leif K. Sandal

SNF Project No. 5183
Managing Risk in Climate – A Dynamic Perspective

The project is financed by the Research Council of Norway

THE INSTITUTE FOR RESEARCH IN ECONOMICS AND BUSINESS ADMINISTRATION
BERGEN 2011

ISBN 978-82-491-0781-0 Printed version
ISBN978-82-491-0782-7 Online version
ISSN 0803-4036

Content

1 Introduction

The motivation behind this article is the hypothesis that postponing action against climate change will have as a consequence that action has to be more pronounced when it is eventually implemented. It is assumed that infinite postponement is simply not an option. By action is primarily meant a cap on total emissions. This cap can be enforced either as an emissions tax or through a scheme of tradeable quotas. This implies that there will be trade-offs between what we save by not paying taxes and the higher consumption of fossil fuel products during the period of postponement on one hand, and the damaged caused by increased aggregate emissions plus the higher tax we will eventually have to pay, on the other hand. In particular it will be interesting to see how much higher an emission tax will have to be for each additional period of postponement and business as usual. In other words, starting early on climate change mitigation results in lower carbon price in the future than postponing mitigation.

In this article a dynamic optimization model in continuous time is applied. An optimal carbon tax beginning in year 2000 is calculated and compared with an optimal tax starting in each consecutive year up to 2070. As the applied model is a closed loop (feedback) model, it is assumed that if the tax is not implemented in 2000, then the aggregated GHG level will increase according to business-as-usual until a tax eventually is implemented. The year when it finally is implemented the GHG level will necessarily be higher than if a tax had been implemented in 2000, and the initial optimal tax will therefore also be higher compared to the optimal path of a tax implemented in 2000. The longer implementation is postponed the higher the initial tax must be. This may result in one of two things, either the initial tax becomes so high that it is politically impossible to implement or the tax that is finally implemented is far too low and hence suboptimal. This way the difference between the two curves, one showing the optimal tax following a scheme implemented in 2000 and the other showing the optimal initial tax implemented in some later year, measures the cost of postponing implementation. This cost will necessarily increase over time as there becomes more catching-up to do.

Given various assumptions about demand and supply of fossil fuel, decay of GHG in the atmosphere and disutility associated with GHG concentration, a model designed to calculate the optimal tax under uncertainty is developed. The latter point is important as the various geophysical processes involved are highly stochastic, in particular the natural decay of GHG over time.

The higher this difference is, the less likely it is that mitigation will be undertaken. In other words, although we may save ourselves many years of not incurring climate change mitigation, it increases the likelihood of future inaction (hysteresis) because it raises mitigation costs in the future to achieve a given concentration target. Thus the longer we wait the greater the chance that the action will be insufficient. Further, it is also our aim to show how much a future tax will have to increase, year by year, for each year that it is not implemented. This will also illustrate the additional cost of each failure of finding an agreement at present and future climate summits.

The present study is in many ways similar to Keller et al. (2007) who made an attempt at putting value on the regrets of procrastination in climate policy. Whereas economic models recommend a reduction in the flow of emissions of about ten to twenty per cent current abatement is only a few per cent and therefore highly suboptimal. Over time the economic cost due to this suboptimal policy may amount to trillions of dollars.

Yohe et al. (2004) take closer look at the argument that near-term action is unwise as it imposes immediate costs whereas the benefits are uncertain and long-term. They take the view that action will sooner or later be necessary and therefore the wait-and-see attitude may not be the optimal one. They ask the question whether we should intervene now as a hedge against the expected cost of meeting an unknown future target. The conclusion is that not only is hedging wise compared to the wait-and-see alternative, it is only a relatively cheap and robust insurance against extreme future events. According to them the best alternative is to introduce an immediate tax at 10 $ per tonne carbon increasing at the rate of interest. Similar results are derived here, but the magnitude of the optimal tax rate and the cost of postponing implementation depends heavily upon the assumptions made about various economic parameters.

2 The conceptual model

In this section the basic model is described. The notation is as follows: x is production of fossil fuel, which also is equivalent to the emission rate as production is measured in emission units, a is concentration of GHG in the atmosphere and τ is the associated ad valorem carbon tax. Concentration a is measured in p.p.m. and emission x is measured in p.p.m. per time unit. The optimal tax is derived starting with the basic variables supply and demand, and then maximizing welfare over time. Welfare is defined as the sum of consumers' and producers' surplus corrected for externalities. The demand function for fossil fuel is denoted by $p(x)$, and the supply function, derived from marginal cost, denoted by $c(x)$.

Welfare, defined as the flow of utility, U, minus the running damage, D, caused by accumulated GHG emissions is given by:

$$W = \int_0^\infty e^{-\delta t} \left[U(x) - D(a) \right] dt$$

where $a(t)$ and $x(t)$ are themselves functions of time. Time, t, is usually suppressed as an argument in the equations for simplicity. Utility, U, defined as the sum of consumer and producer surplus, is formulated a follows:

$$U(x) = \int_0^x [p(y) - c(y)]\, dy$$

where y is used as a dummy variable. Damage, D, is a convex function given by:

$$D(a) = \gamma \cdot ([a - a_d]^+)^2 \tag{1}$$

where $[x]^+ \equiv \max(0, x)$. Here a_d represents the level of GHG below which no damage can be detected.

Demand is assumed to be given by a non-linear function with variable elasticity:

$$p(y) = p_0 \cdot e^{-p_1 \cdot y}.$$

The parameter p_0 represents the maximum price; prices higher than this choke all demand whereas the parameter p_1 is an elasticity parameter. The demand elasticity is given by

$$El_D(y) = -\frac{1}{p_1 \cdot y}.$$

The supply function, on the other hand, is assumed to be linear:

$$c(y) = c_0 + c_1 \cdot y$$

where c_0 and c_1 are supply parameters. The supply elasticity is given by

$$El_S(y) = 1 + \frac{c_0}{c_1 \cdot y}.$$

The elasticities will be used to calibrate the demand- and supply parameters.

This specification of demand and supply implies that the utility flow can be written explicitly:

$$U(x) = \frac{p_0}{p_1}(1 - e^{-p_1 x}) - \frac{1}{2}c_1 x^2 - c_0 x. \tag{2}$$

The ad valorem tax is the difference between the consumer- and producer-price divided by the producer-price:

$$\tau(x) = \frac{p_0 \cdot e^{-p_1 x}}{c_0 + c_1 x} - 1. \tag{3}$$

The equation (3) is not explicitly solvable with respect to x, as it would have been if, e.g., supply and demand had been linear. However, as the dynamic optimization problem is solved numerically, this is of minor importance. The dynamic optimization in this model is subject to a dynamic constraint governing the change in aggregated emissions. The change in aggregated emissions is determined by the flow of emissions subtracted natural assimilation. As natural assimilation is subject to uncertainty, it is modelled as a stochastic process. Production is measured in emission units, and hence the two are equivalent. The dynamic constraint can then be written

$$da = [x - f(a)]\, dt + \sigma dB \tag{4}$$

where $f(a)$ is the deterministic term in the assimilation (or natural adjustment) process. The term da is is an incremental change in the CO_2 concentration level, and x the emission rate as earlier. The term σdB represents an increment of a Brownian motion with variance σ^2. Further it is assumed that the deterministic term, $f(a)$, can be specified as

$$f(a) = \begin{cases} 0 & if \quad a \leq a_0 \\ f_1 - \frac{f_1}{1+f_2(a-a_0)^2} & if \quad a > a_0 \end{cases} . \tag{5}$$

Here f_1 and f_2 are parameters to be estimated and a_0 is by definition the preindustrial CO_2 level as this is the level to which a will converge without any anthropogenic emissions. Based on the data the preindustrial level is assumed to be 280 p.p.m. If, on the other hand, anthropogenic emissions go to infinity $f(a)$ converges to f_1 and $f(a) < f_1$ for all a. Further, the derivative $f'(a) > 0$ for $a \geq a_0$. The parameters f_1 and f_2 are estimated using the ensemble Kalman filter, a method described in a separate section.

2.1 The numerical model: Specification.

In this section the numerical specification of the model is outlined, and each input-function is described step by step. All units are measured in p.p.m. whether it is accumulated GHG, consumption, production or emissions. Present refers to 2007 and the applied discount rate is 2 per cent.

2.2 Demand and supply

The numerical parameters in the demand- and supply functions are specified as follows. The price is normalized such that the present market equilibrium yields a price equal to one, and the present consumption and production are equal to 3.9 p.p.m.. These assumptions combined with assumptions about long-run demand- and supply elasticities yield the parameters in the demand- and supply functions.

The assumption about the demand elasticity is taken from Flood et al. (2010): "While a range of estimates is found, the consensus is that the long-run price elasticity of demand is around - 0.8,...". Regarding supply elasticities of fossil fuel the range is even wider, ranging from almost zero to highly positive depend on which product is in question and whether emphasis is put on short or long term. And there is less consensus. Supply is naturally more elastic in the long run than in the short run, and given the wide range of estimates to choose from in the literature, a long-run supply elasticity of about 1.65 is chosen here as this seems to fall in the middle of this wide range. This is e.g. in accordance with e.g. Dahl & Duggan (1996), but sensitivity analysis with respect to this parameter will be performed later. These assumptions yield the parameters reported in Table 1.

Table 1. *Parameters in the supply and demand functions*

Parameter	Value
p_0	3.5
p_1	0.32
c_0	0.39
c_1	0.1525

2.3 Damage function

The damage function represents a mapping between the aggregated level of GHG and current damage measured in monetary units. This relationship is fairly vague and associated with a high degree of uncertainty. There is some consensus that the annual cost of climate change and global warming in the future probably will range between one and four percent of global gross domestic product (GDP). There is still no consensus that this damage has already started, and therefore parameters are chosen such that total damage varies from 0.2 percent to 4.6 percent of global GDP. Global GDP is calculated as follows: The contribution of fossil energy to the world's domestic product is simply the quantity in the market equilibrium multiplied by the market price. As the market price is normalized to one, and the quantity is 3.9 p.p.m., this contribution is 3.9 in monetary units. Further, fossil energy accounts for approximately 12 - 15 percent of the world's GDP, see BP Statistical Review of World Energy (2010). In the present analysis 12 percent is applied as an estimate implying that the world's gross product is 32.5 measured in the units applied here with 2007 as base year.

With the specification given by (1), it is assumed that an aggregated GHG level below a_d corresponds to zero damage. Here a_d is chosen somewhat arbitrarily to be 350, which was the level around 1960. Implicitly it is assumed that there were no trace of damage associated with GHG prior to 1960. Various values for γ and the corresponding share of GWP (Gross World Product) that the damage accounts for are reported in Table 2.

Table 2. *Parameters in the damage function.*

Case	γ	% of GWP in 2007
1	$4.4 \cdot 10^{-5}$	0.2 %
2	$2.64 \cdot 10^{-4}$	0.9 %
3	$7.92 \cdot 10^{-4}$	2.8 %
4	$9.69 \cdot 10^{-4}$	3.4 %
5	$1.3 \cdot 10^{-3}$	4.6 %
$a_d =$	350	

.

2.4 Natural assimilation

The natural assimilation or decay is a very important function in the model, but at the same time characterized by a high degree of uncertainty. The way it has been formulated here, see eq. (5), it is assumed that a_0 represents the level of aggregated GHG corresponding to the preindustrial level. In other words, this is the natural steady state level of GHG without any anthropogenic emissions. The parameter values in the assimilation function are:

Table 3. *Parameters in the natural decay function.*

Parameter	Value
f_1	2.8797
f_2	$1.378 \cdot 10^{-4}$
σ_0^2	0.0428
a_0	280

3 Carbon data assimilation

The ensemble Kalman filter is used to estimate model parameters of the carbon assimilation function (sometimes called the carbon decay function). A range of different specifications were tried; most had the following characteristics. No (net) assimilation at the preindustrial concentration level (280 p.p.m.). Assimilation increases with the concentration level above the preindustrial level, but after some concentration level, assimilation tapers off. The numbers of parameters in the model is reduced as much as possible as there is limited variation in the data.

3.1 Data

The data on CO_2 emissions (1751 – 2007) were collected from the Carbon Dioxide Information Analysis Center at Oak Ridge National Laboratory (Boden & Marland 2008). Unfortunately, observation uncertainty was not provided.

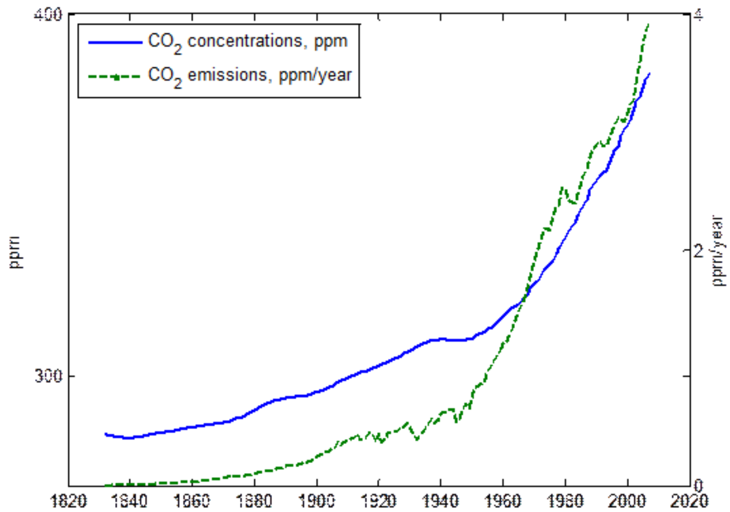

Figure 1: *CO2 concentrations (solid line, left axis) and emissions data (dotted line, right axis).*

The atmospheric concentration of CO_2 has been measured at Mauna Loa, Hawaii since 1959 (see Tans and Keeling 2011). Measurements are made with high accuracy. In the assimilation exercise it is assumed more uncertainty in the concentration data as they are regarded as globally representative.

CO_2 concentrations data prior to 1959 were collected from the Law Dome Atmospheric CO_2 data, that is, observations made from ice core samples (Etheridge et al. 2001, the data was first published in Etheridge et al. 1996). The ice core data contains yearly concentration observations for the years $1832-1978$ and five year averages for the years $1010-1975$. As the yearly observations stretched into the preindustrial period, we chose not to use the long time series. In the years where the Mauna Loa and Law Dome data overlapped, the Mauna Loa data were used. Using the Law Dome data would not make much of a difference as they coincide to a large degree. Data are shown in Figure 1.

3.2 The Model

The carbon assimilation model is the one described by eq.(5). It is readily seen that $f(x) < f_1$ for all a and that $f(a)$ converges to f_1 when a goes to infinity. The first derivative is always positive when $a \geq a_0$. The parameters in equation (5) are estimated with the ensemble Kalman filter. How the concentration level changes is given by eq. (4).

In (4), da is an incremental change in the CO_2 concentration level, x is the emission rate (see Figure 1), dt is the incremental time step, and σdB describes an increment in Brownian motion process with variance σ^2.

3.3 The Ensemble Kalman Filter

The ensemble Kalman filter is used to estimate the carbon assimilation function in equation (5). The filter is a data assimilation method which is widely used in physical applications like meteorology and oceanography, where phenomena typically have a chaotic nature. It has structural relationships to the classical Kalman filter, but extends to a large class of nonlinear models (see Burgers et al. 2008 and references therein). The method was first suggested by Evensen (1994), while Burgers et al. (2008) provided a theoretical clarification. Evensen (2003) reviews both theoretical developments and survey applications of the ensemble Kalman filter and related techniques. Evensen (2009) discuss more recent developments.

The ensemble Kalman filter use a Markov Chain Monte Carlo approach to solve the Fokker-Planck equation which governs the time evolution of models like the model in (4). The method applies to state space models with the dynamic equation (the state or model equation) written as a stochastic differential equation (see equation (4)). In addition, the measurement equation relates observations to model states. In many applications, ours included, the observed phenomena is observed directly and the measurement equation becomes trivial:

$$z = Ma + \varepsilon \tag{6}$$

where z is the observed concentration, a is the modeled concentration, and ε is a normal distributed error term with known variance. M is called the observation or measurement operator and is in our case simply the identity operator. The filter procedure works as follows: An ensemble of model states; a cloud of points in the state space, is integrated forward in time according to the stochastic model dynamics. Error terms are simulated. Each point in the forward integrated cloud represents model forecasts. At observation times, the Kalman gain is calculated using ensemble moments instead of the unknown true moments. The integrated ensemble is then updated with the gain matrix which defines how much weight to assign to observations and to forecasts. It can be shown that the mean of the updated ensemble converges to the classical Kalman filter solution asymptotically in the ensemble size (Evensen 2009). Usually, a limited size of the ensemble is sufficient for reasonable statistical convergence. The ensemble size necessary to represent the density will depend on the application. If the ensemble mean is treated as the best estimate, it can further be shown that the second moment of the ensemble is the best estimator of the error covariance (Burgers et al. 1998).

The ensemble Kalman filter can be used for parameter estimation. Unknown parameters are simply added to the state-space and treated as model states (Evensen 2009, p. 101). When unknown parameters are present, some adjustments of the observation operator and the Kalman gain matrix are necessary, see Evensen (2009) for details. As for all Kalman filter methods, one must define initial conditions; priors in a Bayesian framework. With unknown parameters, initial conditions can be difficult to specify, but usually one has theory or earlier work to rely on.

When ensemble members are integrated forward in time, one can in principle use any forward integration method, like a Runge-Kutta or Euler method. Here, however, equation (4) is discretized by setting $dt = \frac{1}{10}$ (measured in years), and the discretized equation is used directly. Errors are simulated. The method requires observations to be treated properly as random variables, which in a Monte Carlo approach implies that measurement errors in equation (6) have to be simulated as well; once for each observation for each state ensemble member (Burgers et al. 1998). The approximation introduced with the ensemble representation of observation errors does not influence the state estimate and can be made less than the true observation uncertainty with a large enough ensemble size (Evensen 2009, pp. 41-42). Let ψ denote the state vector extended with the unknown parameters; Ψ denotes the ensemble of states. ψ^f denotes the forecasted state vector, similarly for the ensemble matrix; Ψ^f. The Kalman gain matrix is given by

$$K = (C_{\Psi\Psi})^f M^T (M(C_{\Psi\Psi})^f M^T + C_{EE})^{-1} \tag{7}$$

where $(C_{\Psi\Psi})^f = cov(\Psi^f, \Psi^f)$ is the covariance in the ensemble of forecasted states, M^T is the transpose of the measurement operator, and C_{EE} denotes the covariance in the ensemble of simulated measurement errors E. The expression for the Kalman gain is similar to the expression in the standard Kalman filter, but the generally unknown covariances are replaced by ensemble moments. The forecasted ensemble member j is updated in what is known in the technical literature as the analysis step, where the Kalman gain acts as a weight between the forecasted state and the observation:

$$\overline{\Psi(j)} = \Psi(j)^f + K(d + E(j) - M\Psi(j)^f). \tag{8}$$

With the identity operator as the measurement operator, one readily sees that with $K = 0$, no weight is given to the perturbed observation $d+E(j)$. With $K = 1$, no weight is given to the forecast. The mean of $\overline{\Psi}$ is then the ensemble Kalman filter estimate of the state vector contingent on the observation d. Usually, one has a time series of observations; the ensemble of states is then forecasted until the next observation time, a new gain matrix is calculated, and the ensemble is updated in a new analysis step. The sequential nature of the method improves efficiency of the numerical implementation. It also means that estimates are conditional only on observations up until a given time. If estimates conditioned upon all available observations are desired, estimates can be smoothed (Evensen 2009).

3.4 Results

The analysis is done for the years where there are observations of both emissions and concentration levels; 1832 – 2007. Parameter f_2 is assumed lognormal distributed (thus nonnegative). Initial conditions, used to generate the randomly drawn initial ensemble, are given in Table 4. Further, 500 ensemble members are used, and observation uncertainty (variance of error term in equation (6)) is 0.12 (see Tans and Keeling 2011).

Table 4. *Initial conditions.*

	mean	st. d.
x	284.3	10
p_1	3	1
$\ln p_2$	-9.21	3

Figure 2 shows the estimated concentration level and demonstrates that the

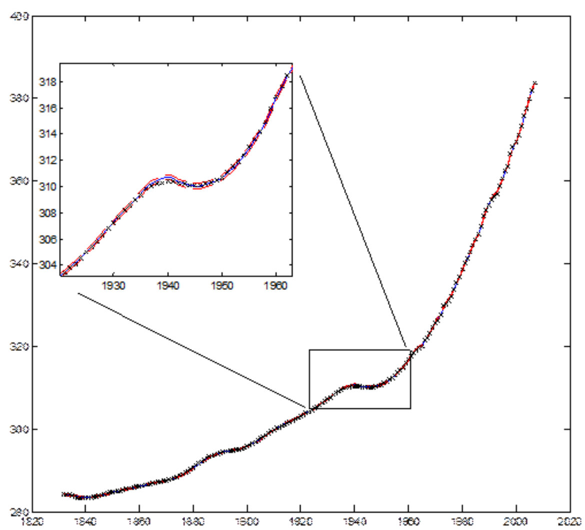

Figure 2: *Estimated carbon dioxide atmospheric concentration level. Dashed lines show 95% confidence limits.*

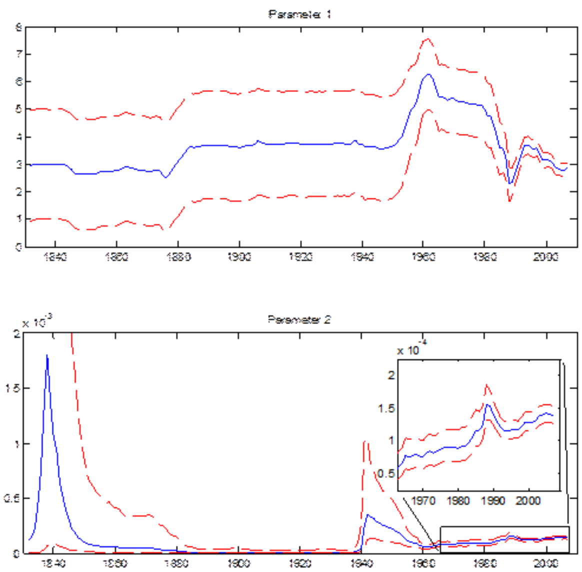

Figure 3: *Parameter estimates, see equation (1). Dashed lines show 95% confidence limits.*

model to a large extent explains the observed levels. The figure also shows confidence limits of the estimate, and they are quite narrow as a result of the small observation uncertainty. It could be argued that while the Mauna Loa observations are highly precise, uncertainty should be added when they are treated as representative of the global concentration level. However, it is not obvious how much uncertainty to add, and it is therefore left for future research.

Figure 3 shows parameter estimates. The final estimates, which are contingent on all data, are $f_1 = 2.88$, and $f_2 = 1.38 \cdot 10^{-4}$. The parameters are highly significant, but it is worth noting that if the sample had lasted only to around 1960, the estimate for f_1 would be much higher. As f_1 is the upper bound for the yearly assimilation rate of carbon, it is an important parameter when forecasting emissions and discussing mitigation. Thus, if one wants to be optimistic when it comes to the potential for natural cleaning of atmospheric carbon dioxide, and at the same time is willing to overlook the last 50 years of observations, one could argue that the parameter should be around 6.2.

The size of the noise term, governed by σ, in equation (4) is also estimated. The estimate is shown in Figure 4. The final estimate is at $7.7{\cdot}10^{-3}$. The estimate is, however, biased downward. A more appropriate estimate is shown in Figure 5. The estimate is very volatile, and, further, only an estimate of the noise in each period. The nature of the noise process makes the estimate valid only locally in time. As seen in equation (4), we are interested in a time-independent estimate. Here the average of the estimate shown in Figure 5 is used, which is 0.0428. Another alternative would be to use a moving average, but the size of the window would matter. A ten year window at least seems to capture the main features of the yearly estimate, see Figure 5. The average for the last ten year window is 0.0764. The time structure of the curves in Figure 5 is interesting in itself. Historically, two periods had increased noise in the atmospheric carbon dioxide concentration; the periods around 1880 and 1940. Both periods corresponds to when concentration levels started to increase faster than before (see Figure 1) and to when parameter estimates changed (see Figure 3). After 1960, the noise has increased steadily; again, it seems that concentration levels has increased faster and faster, and parameter estimates are less stable.

With the interpretation of the Kalman gain as a weight in equation (8), it could be of interest to study the gain itself. Figure 6 shows the Kalman gain for all state variables. It can be shown that with the observation equation (6), where the observation operator M is the identity, the gain is bounded by zero and one. Most of the time, the gain for the carbon dioxide concentration level is slightly above 0.5; the estimate is thus approximately the average of

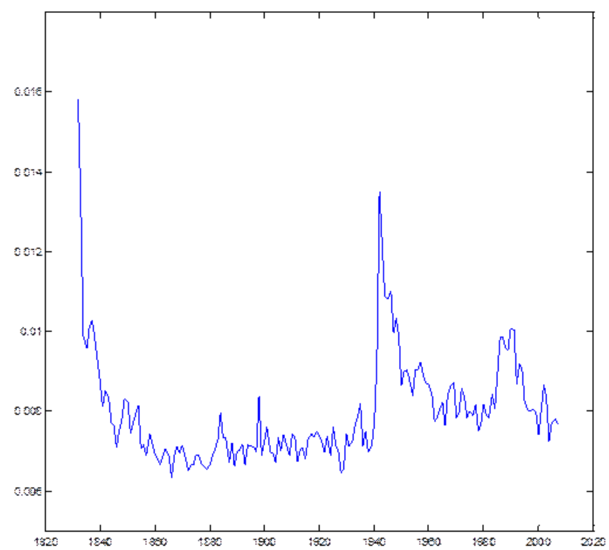

Figure 4: *Estimated noise level, see equation (4).*

the forecasted estimate and the observed state. The bounds on the gain do not hold for parameter states, however. For parameter states, the observation operator is the zero operator. Thus, the Kalman gain scales the influence of the observed state on the estimated parameter. The influence can be both positive and negative, as seen in Figure 6. Compared to the events observed in Figure 5, there seems to be little connection to the Kalman gain apart from generally larger influences approximately after 1960.

Finally, in Figure 7 the estimated assimilation function is presented, see equation (5). Missing emissions, emissions minus the yearly increase in atmospheric concentration, are also plotted. The estimated function seems to be a reasonable approximation, but possibly overestimated at low concentrations.

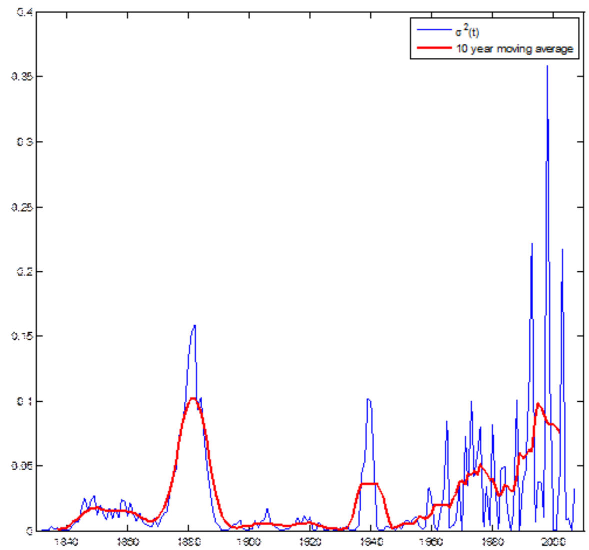

Figure 5: *Model error.*

3.5 Alternative Model

An alternative model, which links the concentration of carbon dioxide to the global average temperature, is also developed and estimated. The idea is that damages might be closer linked to increases in temperature rather than the level of carbon dioxide concentrations.

For the alternative model, a slight variation to the carbon assimilation function discussed above is chosen, see equation (4):

$$f(a) = \begin{cases} 0 & if \quad a \leq a_0 \\ f_1 - \frac{f_1}{1+f_2(a-a_0)^{\frac{3}{2}}} & if \quad a > a_0 \end{cases} \tag{9}$$

The dynamic model consists of equation (4) in addition to an equation which links temperature (b) to the carbon dioxide concentration:

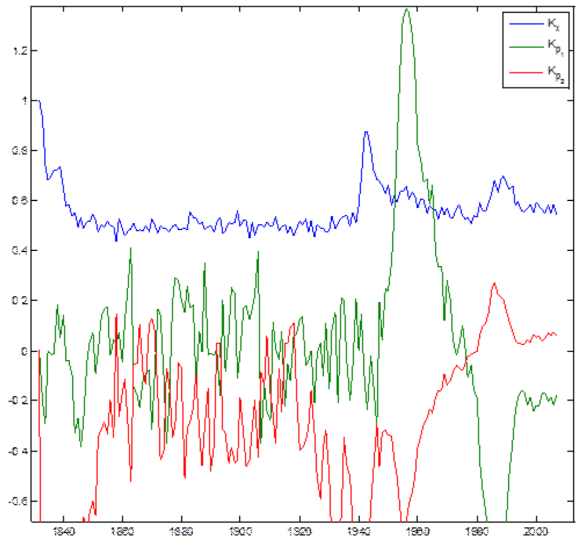

Figure 6: *Kalman gain for all state variable*

$$
\begin{aligned}
da &= [x - f(a)]\, dt + \sigma_a dB_a \\
db &= \left[\frac{f_3}{\ln 2}\frac{1}{a}\right] da + \sigma_b dB_b
\end{aligned}
\tag{10}
$$

The system (10) is inspired by Scheffer et al. (2011) and references therein. The interpretation of the parameter f_3 is relatively simple; it measures the increase in the mean global average temperature with double the carbon dioxide concentration. See Scheffer et al. (2011) for further discussion of the carbon-temperature model.

Data on the global average temperature was collected from National Aeronautics and Space Administration (2011), which provides updates to Hansen et al. (2006). Temperature data was available from 1881 to 2010. Note that

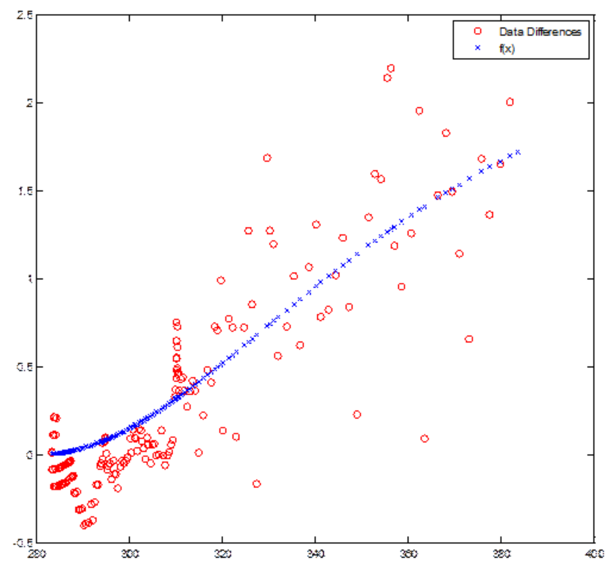

Figure 7: *Estimated assimilation and data differences.*

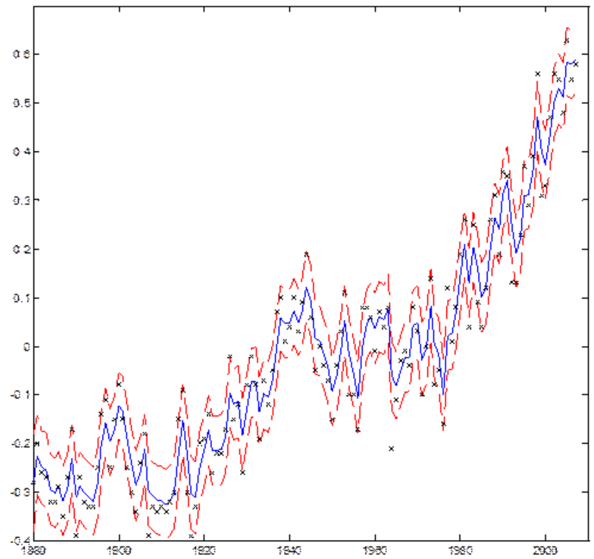

Figure 8: *Estimated global average temperature anomaly (blue curve). Dashed curves show 95% confidence intervals, while x-marks show the observed anomalies.*

temperature is measured as anomalies relative to the average temperature in 1951 – 1980; see Hansen et al. (2006) for details.

Figure 8 shows the estimated temperature anomaly. The estimated carbon dioxide concentration is as least as good as the estimate in Figure 2. Figure 9 shows parameter estimates.

The results from the alternative model holds promise for further research. A challenge is then to utilize (10) in an integrated economic model.

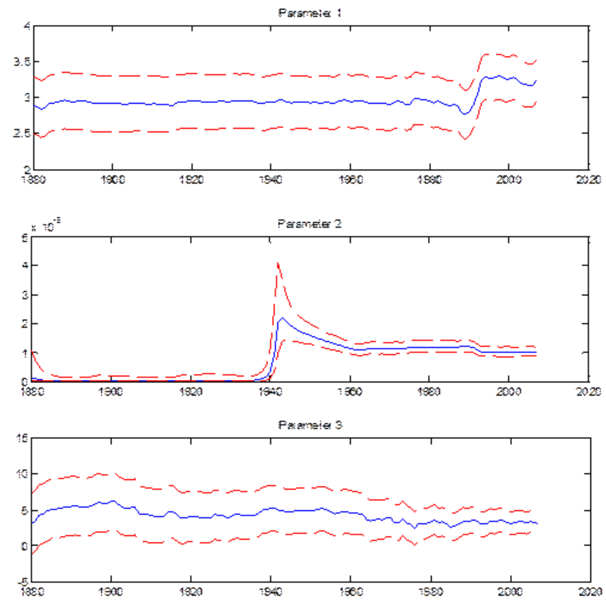

Figure 9: *Parameter estimates, see equations (6) and (7). Dashed curves show 95% confidence limits.*

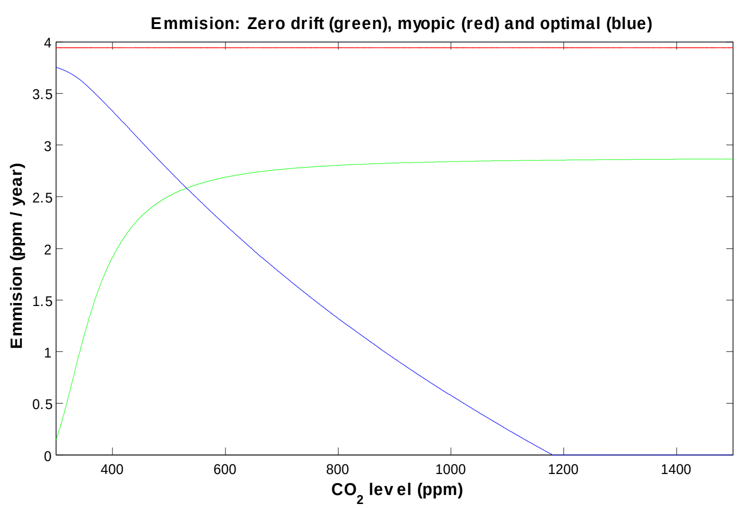

Figure 10: *Damage equal to 0.2 % of GWP in the initial year 2000.*

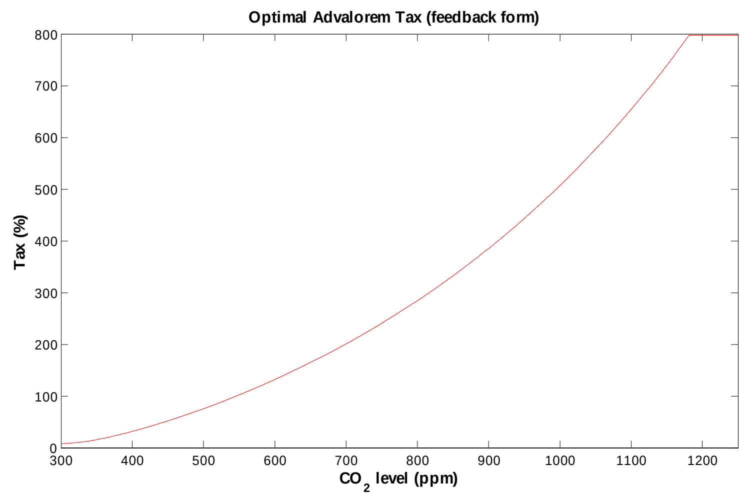

Figure 11: *Damage equal to 0.2 % of GWP in the initial year 2000.*

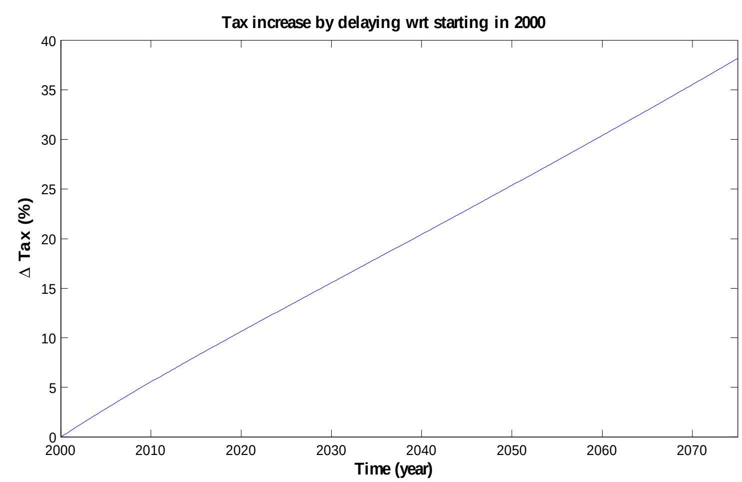

Figure 12: *Damage equal to 0.2 % of GWP in the initial year 2000.*

4 Model results

In this section some results based on this model are presented. The cases outlined here are the five cases suggested in Table 2 where the damage coefficient, γ, is the only factor that varies. The variation in γ corresponds to a variation of the damage relative to GWP from 0.2 per cent to 4.6 per cent, in other words a fairly large range of possible outcomes. This parameter is regarded as, perhaps, the most uncertain parameter in the model, and hence the large variation.

Case 1 is depicted in Figures 10, 11 and 12. In this case the damage due to climate change accounts for only 0.2 per cent of GWP at the current level. Figure 10 illustrates emission as a function of the aggregated CO_2- level. Business as usual will be around 3.9 and increasing (red curve). The downward-sloping blue curve is the optimal emission level and the green curve represents natural assimilation with zero drift. The intersection between the blue and green would then be the long-term steady state if the model had been deterministic. This level would then be 536 p.p.m. with emissions equal to 2.6 p.p.m..

Figure 11 illustrates the optimal tax against aggregated CO_2 - level. It starts at some 10 per cent and hits the maximum level 800 per cent at an aggregated GHG level of almost 1200 p.p.m.. The reason why 800 per cent tax is a "ceiling" is that this tax level will curb all demand and is therefore the highest tax level needed. Figure 12 shows the increment in the optimal ad valorem tax for each year implementation of the tax is postponed compared to if it was implemented in year 2000 right after the first Kyoto-meeting. If, e.g. implementation is postponed until 2070 it must be more than 35 per cent higher than the one implemented in 2000. This corresponds to an increase of about 0.5 per cent per year. This may not sound too scary, but it is calculated for a case with almost negligible damage caused by global warming, in other words the most optimistic case.

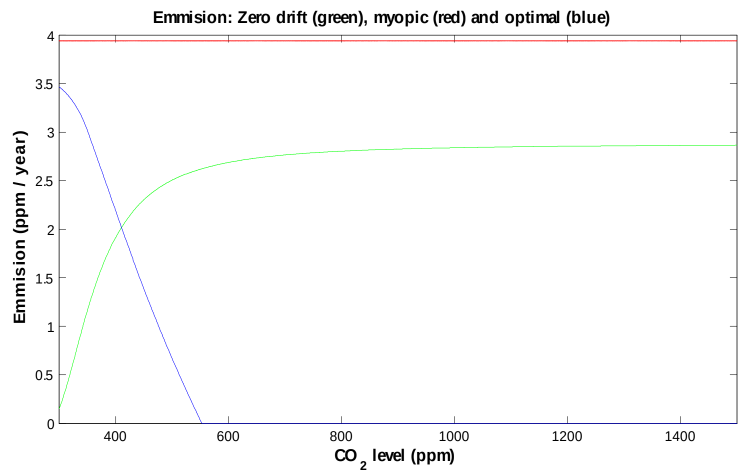

Figure 13: *Damage equal to 0.9 % of GWP in the initial year 2000.*

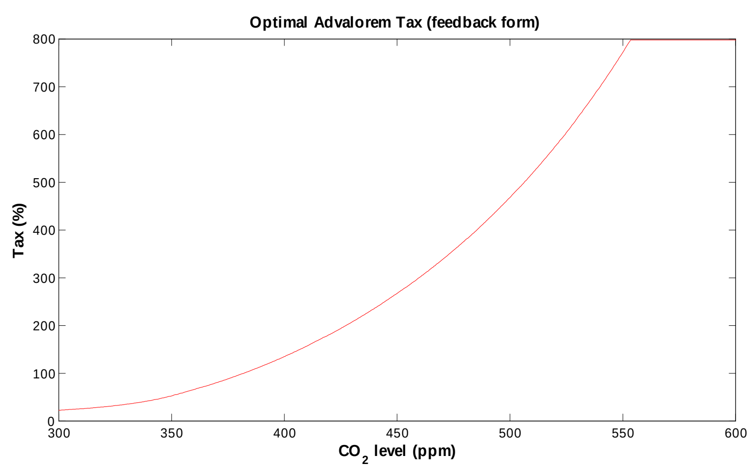

Figure 14: *Damage equal to 0.9 % of GWP in the initial year 2000.*

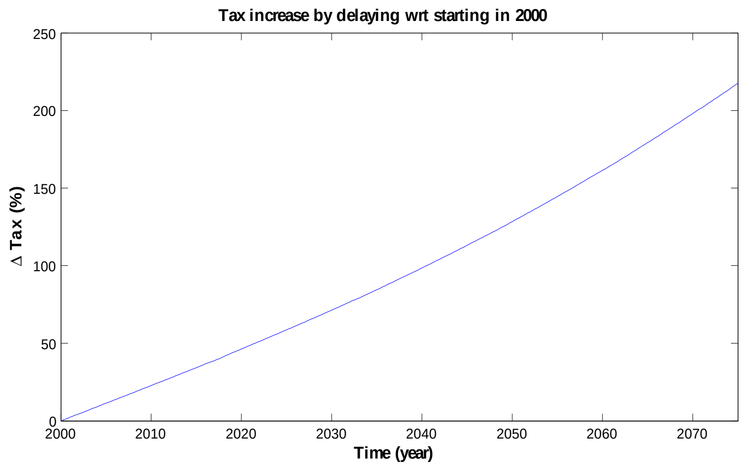

Figure 15: *Damage equal to 0.9 % of GWP in the initial year 2000.*

Figures 13, 14 and 15 depict case 2 where the damage accounts for 0.9 per cent of world GDP at the current level. Figure 13 corresponds to Figure 10, Figure 14 corresponds to Figure 11, etc. The deterministic steady state would be 412 p.p.m. with an emission equal to 2 p.p.m.. From Figure 14 it is seen that the optimal tax hits the "ceiling" already at about 550 p.p.m.. Figure 15 shows the increase in the tax implied by postponing implementation. Here it is seen that by waiting 70 years from year 2000, the initial tax must be increased by more than 200 per cent or around three per cent per year.

In Figures 16, 17 and 18 Case 3 is depicted where the damage accounts for 2.8 per cent of GWP at current levels. The figures appear in the same order as earlier. In this case it would be optimal to stabilize GHG concentration at about 380 p.p.m. which is close to the present level and only emit 1.7 p.p.m. per year from now on. This would require an ad valorem tax starting at more than 200 percent as of year 2000 and increasing fairly rapidly for each year it is postponed. As seen from Figure 18, if implementation of the tax is postponed until the late 2030s for example, the initial tax must be around 275 per cent

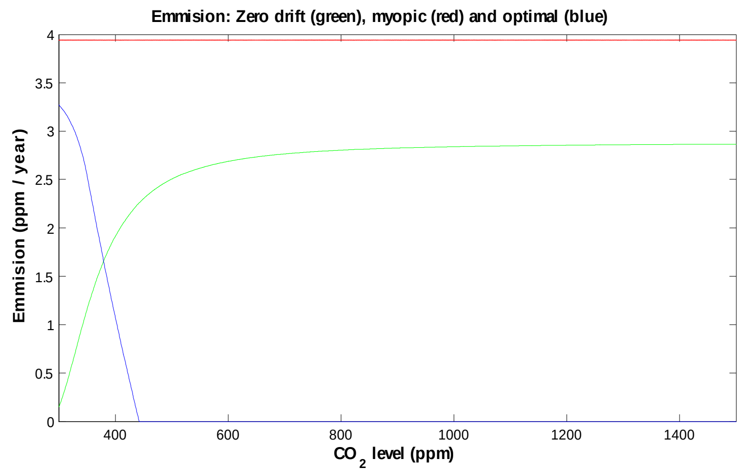

Figure 16: *Damage equal to 2.8 % of GWP in the initial year 2000.*

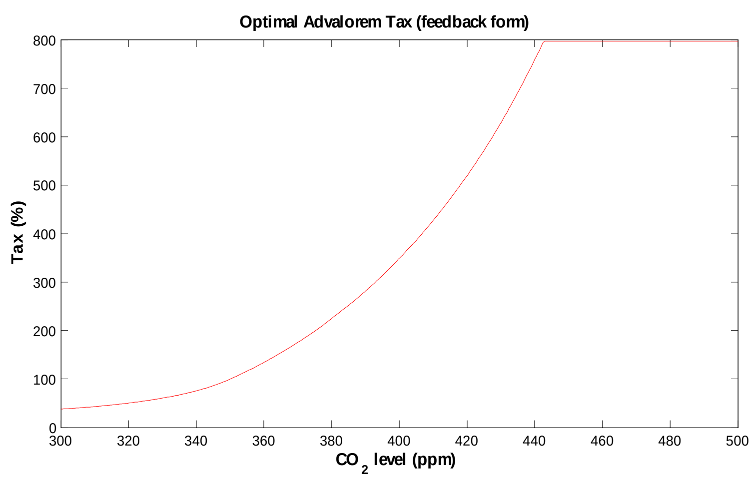

Figure 17: *Damage equal to 2.8 % of GWP in the initial year 2000.*

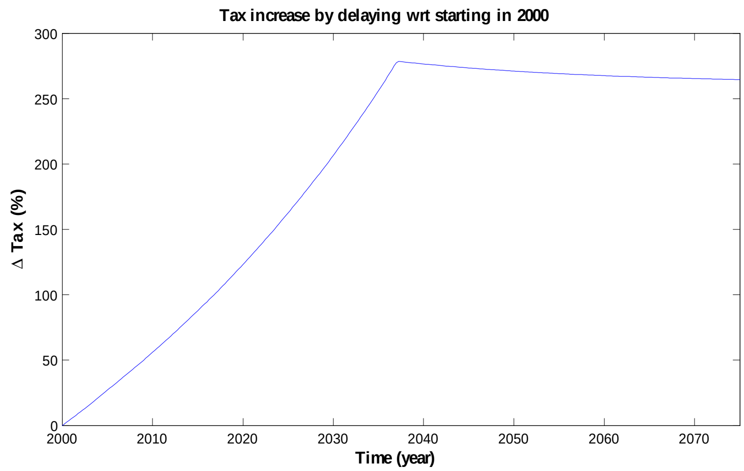

Figure 18: *Damage equal to 2.8 % of GWP in the initial year 2000.*

higher than a tax starting in year 2000 equivalent to an increase of about 7-8 per cent per year. The reason why the curve flattens out after some time is that the tax after that must be sufficiently high to choke all emissions. And from Figure 17 it is seen that the optimal tax increases rapidly also with GHG concentration itself and hits the ceiling for a concentration level which is not much higher than 400 p.p.m..

Case 4, illustrated in Figures 19, 20 and 21, represents a case where the current damage relative to GWP is 3.4 per cent. The deterministic steady state goes down to 376 p.p.m. with an emission rate equal to 1.6 p.p.m.. From Figure 21 it is seen that the tax must be increased by 260 per cent if it is postponed until 2030, which on average is 8 - 9 per cent for every year it is postponed.

In Figures 22, 23 and 24 Case 5 is depicted where the current damage is as high as 4.6 per cent of GWP. The deterministic steady state is down to 370 p.p.m. and the corresponding steady state emission 1.5 p.p.m. which is less than 40 per cent of today's actual emissions. The optimal tax hits the ceiling at 400 p.p.m., which is approximately the same as in Case 4. In Case 5 it is

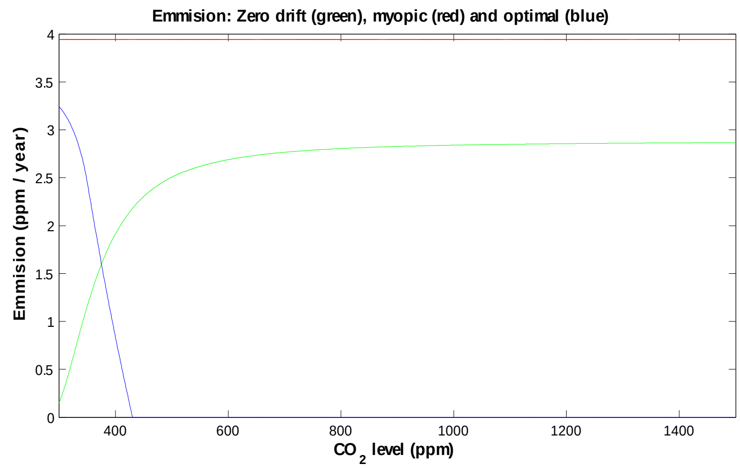

Figure 19: *Damage equal to 3.4 % of GWP in the initial year 2000.*

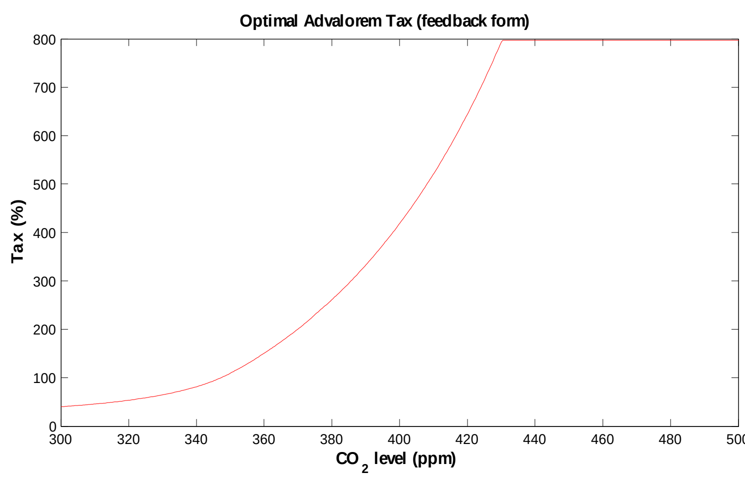

Figure 20: *Damage equal to 3.4 % of GWP in the initial year 2000.*

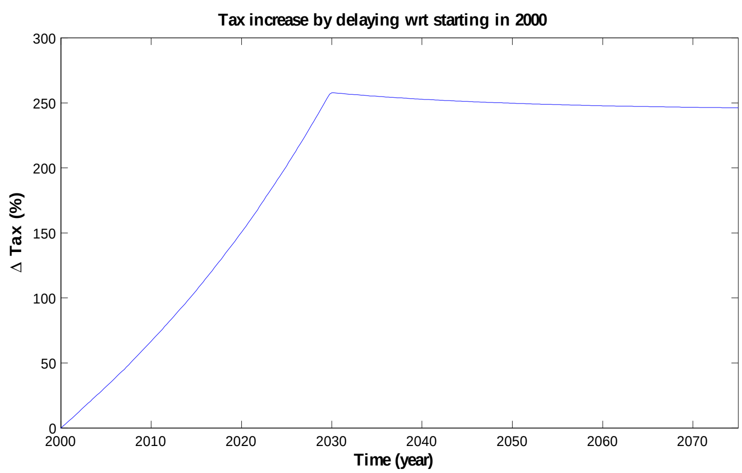

Figure 21: *Damage equal to 3.4 % of GWP in the initial year 2000.*

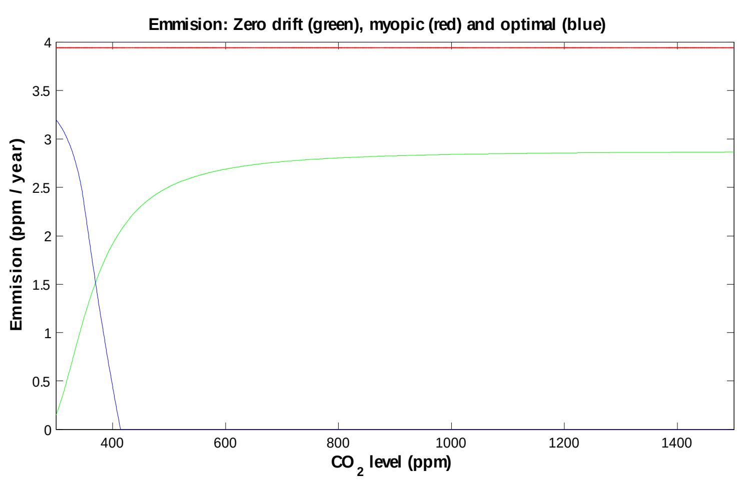

Figure 22: *Damage equal to 4.6 % of GWP in the initial year 2000.*

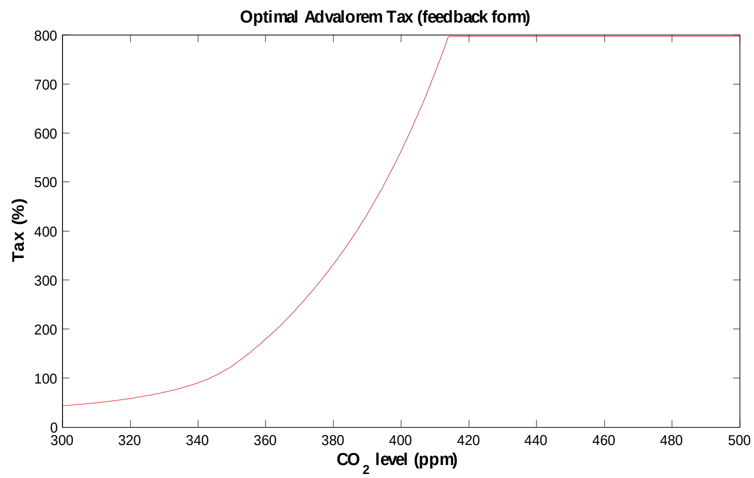

Figure 23: *Damage equal to 4.6 % of GWP in the initial year 2000.*

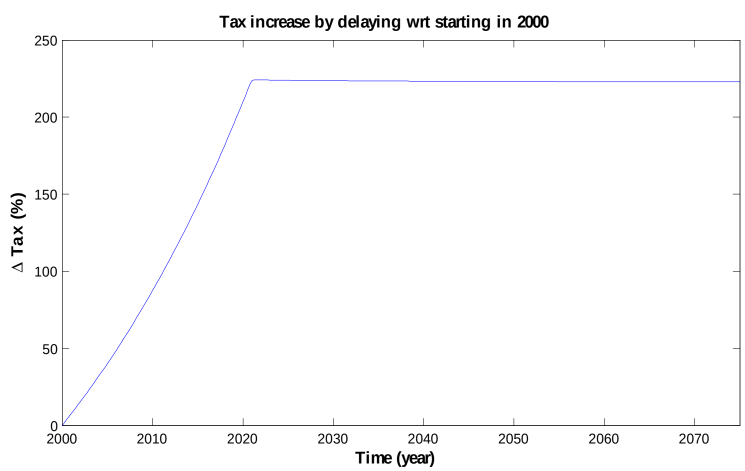

Figure 24: *Damage equal to 4.6 % of GWP in the initial year 2000.*

seen that only by postponing the introduction of a tax further than 2020 means that it must be 230 per cent higher than if it had been introduced in year 2000 or increase by 11 - 12 per cent per year.

5 Sensitivity analysis

In this section sensitivity analysis with respect to supply and demand is performed. The values in the basic model were -0.8 for the demand elasticity and 1.65 for supply elasticity. The value of the supply elasticity is regarded more uncertain than the demand elasticity. The sensitivity analysis is therefore performed as follows: for the demand elasticity it is varied by +/- 0.1 and for the supply elasticity by +/- 0.55. In all cases the intermediate case with respect to damage is applied, that is case 3 in Table 2 where the damage accounts for 2.8 per cent of global gross product in 2007. This means that each of the figures in this section can be compared to one of the figures 16, 17 or 18. The various alternatives and their consequences for optimal tax and emissions will be dealt with one by one.

5.1 Variation in supply elasticity

The results when the supply elasticity is reduced from 1.65 to 1.1 are illustrated in figures 25, 26 and 27. Optimal emissions are not affected by a change in supply elasticity but the optimal tax is heavily affected. It hits the ceiling at about the same level as earlier but the ceiling is much higher. In fact a tax of some 3750 per cent is necessary in order to curb production and consumption. As a consequence the increase in the tax resulting from postponing implementation must also be higher. If, for example, implementation is postponed until 2040, the tax must be 1200 per cent higher than if it was implemented in 2000. This represents an increase in the tax by about 30 per cent for each year it is

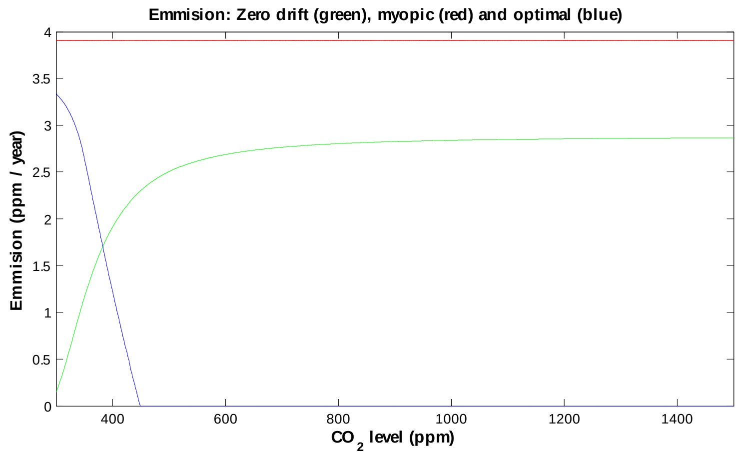

Figure 25: *Supply elasticity equal to 1.1.*

postponed.

The results when the supply elasticity is increased from 1.65 to 2.2 are illustrated in figures 28, 29 and 30. Again optimal emissions are not affected but, just as tax rates increase when supply becomes less elastic, tax rates decrease when supply becomes more elastic. The "ceiling" in Figure 29 is now reduced from 800 per cent to some 550 per cent and from Figure 30 it is seen that by postponing implementation of the tax until 2040 the tax must increase by about 220 per cent compared to 2000, or by about 5 per cent for each year it is postponed. The intuition behind these changes is fairly simple. The more inelastic supply is, the more of the tax burden will be transferred to the producer. The tax rate must be relatively high in order to have any effect on production and the environment. When supply is elastic a much lower tax rate is needed in order to achieve the desired effect which is to internalize the externality. These qualitative results may come as no surprise but the purpose here is to investigate numerically how much tax rates have to change when elasticities are varied. The range of supply elasticities analysed here is quite wide as this is a fairly

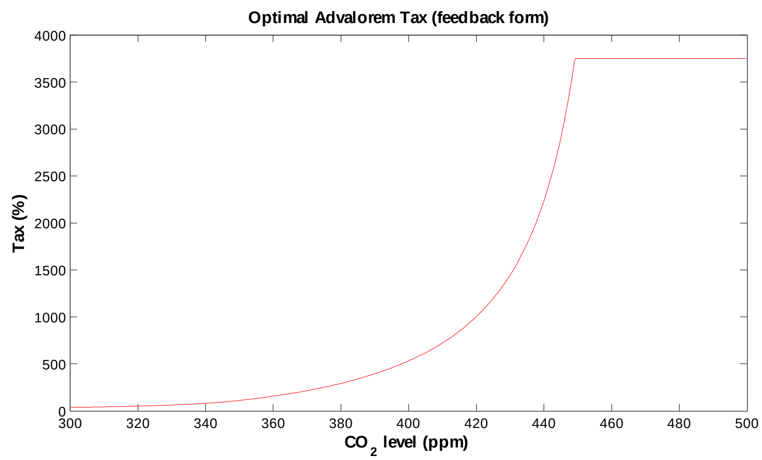

Figure 26: *Supply elasticity equal to 1.1.*

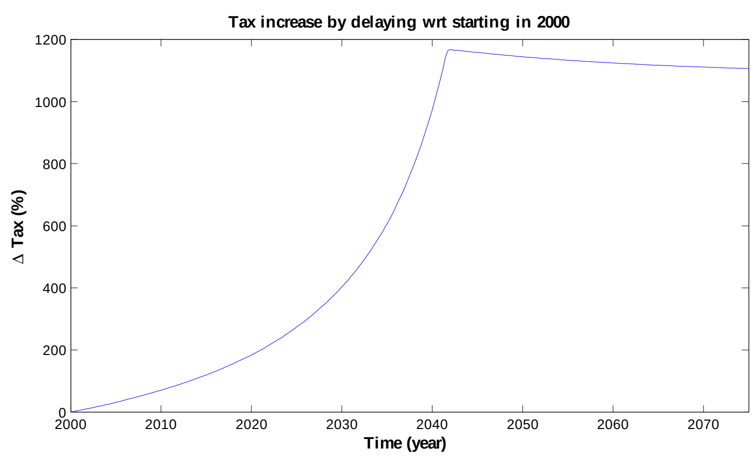

Figure 27: *Supply elasticity equal to 1.1.*

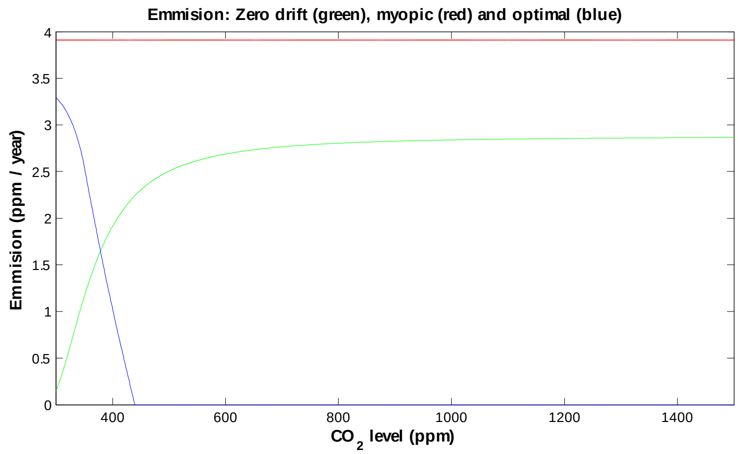

Figure 28: *Supply elasticity equal to 2.2.*

uncertain parameter, and therefore the effect on the tax rates is also quite large. The underlying physical emissions remain the same; it is only tax rates that are affected.

5.2 Variation in demand elasticity

The results when the demand elasticity is reduced (in absolute value) from -0.8 to -0.7 are illustrated in figures 31, 32 and 33. When the demand elasticity is reduced, the same qualitative effects are expected as when the supply elasticity is reduced. But as the reduction in the demand elasticity is less than the reduction in the supply elasticity, the numerical effect is less. The maximum tax rate needed now increases from 800 per cent to almost 1000, and the relative increase in tax caused by postponing implementation of the tax is now 160 per cent of the tax in 2000 is implementation is postponed almost 40 years as compared to the 280 per cent it must increase in the basic model, see Figure 18. This is equivalent to about 4 per cent for each year it is postponed.

The results when the demand elasticity is increased (in absolute value) from

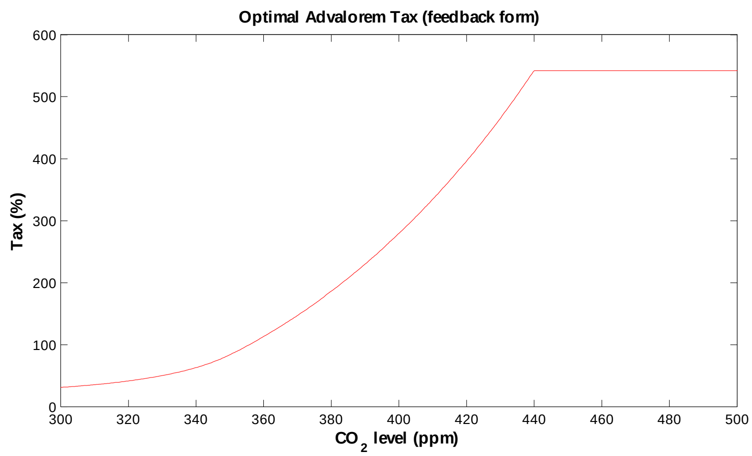

Figure 29: *Supply elasticity equal to 2.2.*

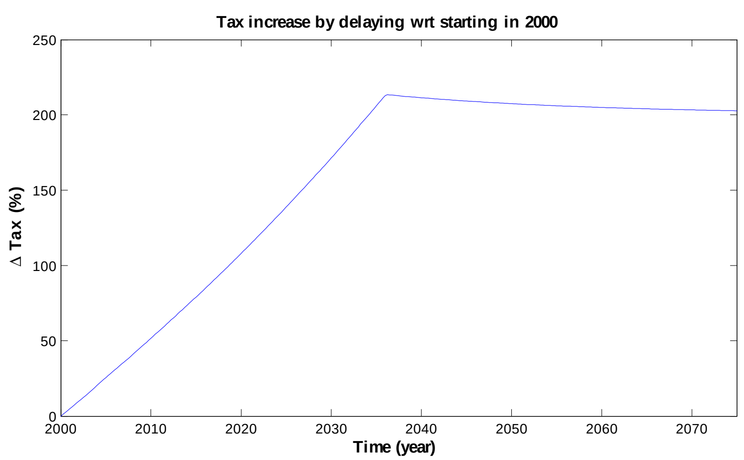

Figure 30: *Supply elasticity equal to 2.2.*

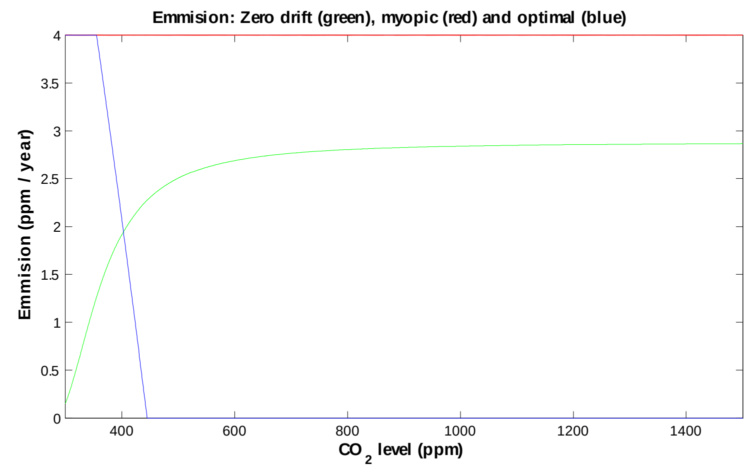

Figure 31: *Demand elasticity equal to -0.7.*

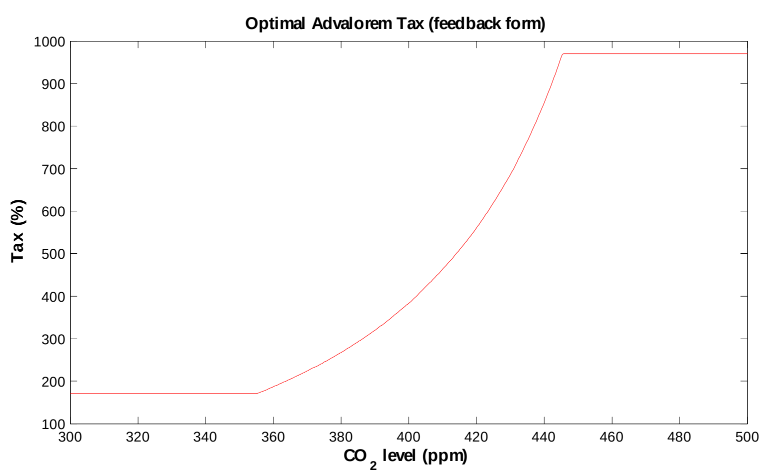

Figure 32: *Demand elasticity equal to -0.7.*

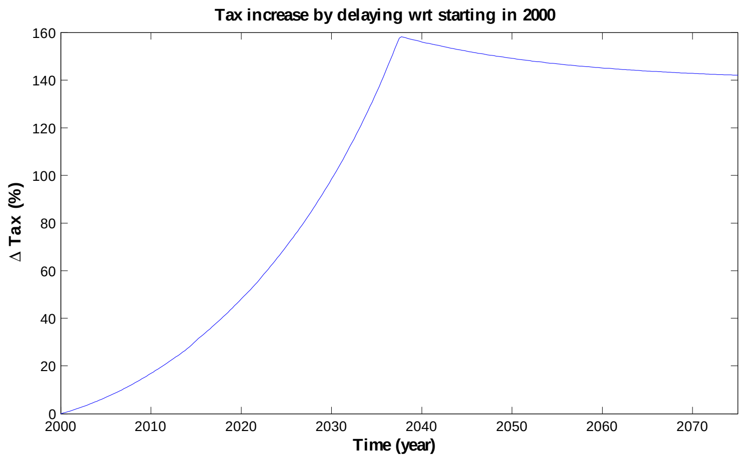

Figure 33: *Demand elasticity equal to -0.7.*

-0.8 to -0.9 are illustrated in figures 34, 35 and 36. As expected tax rates will now decrease but not as much as they decrease when the supply elasticity is varied simply because the variation in supply elasticity was larger. In this case the maximum tax rate, which curbs all production, is just below 700 per cent and a tax rate implemented in year 2040 must be about 150 per cent higher than if the optimal tax had been implemented in year 2000.

6 Summary and conclusions

The main purpose of this report is to investigate the effects of postponing implementation of a carbon tax assuming that externalities associated with climate change and global warming is real and failing to internalize these externalities will only lead to a suboptimal situation. The model applied to investigate these externalities start with the basic relationships, namely supply and demand for fossil fuel, and an added damage term that accounts for the externality. The objective is then to maximize the sum of consumers' and producers' surplus ad-

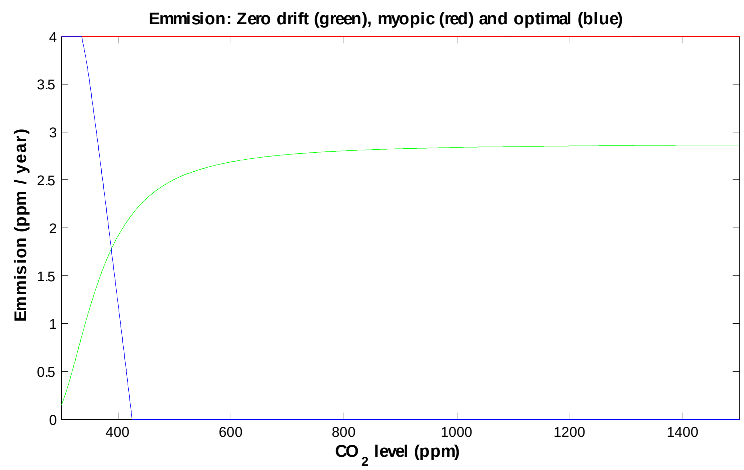

Figure 34: *Demand elasticity equal to -0.9.*

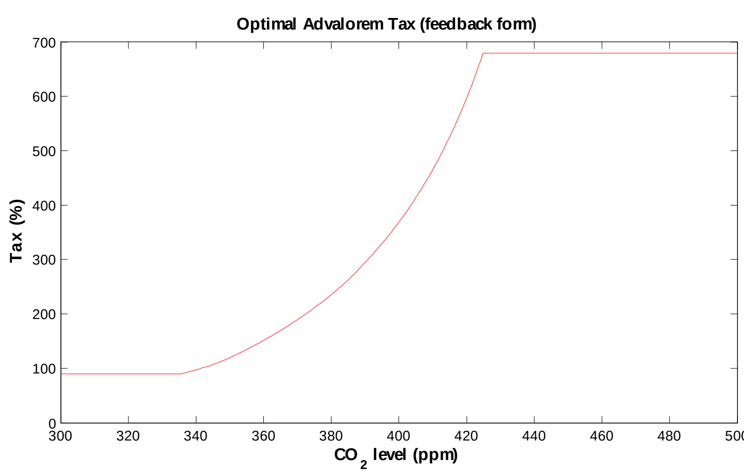

Figure 35: *Demand elasticity equal to -0.9.*

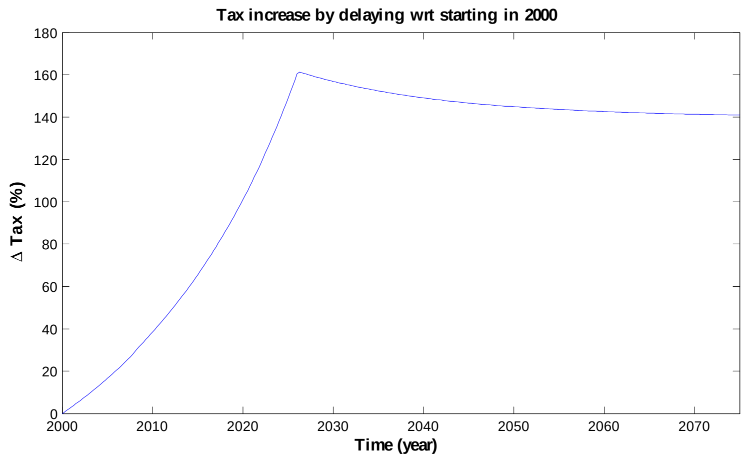

Figure 36: *Demand elasticity equal to -0.9.*

justed for the externality. This must be done subject to the dynamic constraint derived from emissions of carbon associated with extraction and consumption of fossil fuel and the natural assimilation of carbon in the atmosphere. The model is solved as a closed loop feedback policy. First the optimal emission path is calculated, then the corresponding tax path is found. As the externality dealt with here is a pure stock externality it turns out that the optimal tax is equal to the shadow cost of the pollutant.

The dynamic equation for assimilation, or natural decay, of carbon is specified using a fairly sophisticated method, namely the ensemble Kalman filter. Given the relative simplicity of the model with only one type of production and one type of pollutant, this method is supposed to give a best possible estimate of the parameters in the assimilation function.

Regarding the damage term there is still no consensus among scientists about the degree of the damage or the form of the damage function. In this report a convex (quadratic) function is applied, but a large range of parameter values are investigated to test for the sensitivity of the results to variation in the magnitude of the damage. Measured as a percentage of the world's gross product this variation corresponds to a variation from about 0.2 per cent to 4.6 per cent. In other words, from almost nothing to up to almost five per cent. This range ought to cover the beliefs of most scientists from the most optimistic to the most pessimistic ones. Altogether five different cases were analysed. Here the main results from three of them are summarized.

From the basic model it was seen that if the damage parameter is adjusted such that the magnitude of the damage at present only accounts for 0.2 per cent of the world's domestic product (GWP) then it is not necessary to curb all extraction of fossil fuel until the aggregated level of carbon has reached 1200 p.p.m. The optimal tax must increase by only 0.5 percent for each year it is postponed. In other it must 35 per cent if it is implemented in 2070 than in 2000. The optimal steady state in the deterministic counterpart of the model is an aggregated level of just above 500 p.p.m. with an annual emission around 2.5 p.p.m.

In the intermediate case the damage today is supposed to represent 2.8 per cent of GWP today. Emissions must be choked already at an aggregated level around 450 p.p.m., and the same tax rate as earlier, 800 per cent, is needed. This does not change as the underlying economic parameters remain unchanged. The tax must increase by almost 7 per cent for each year it is postponed until it is 275 per cent higher in 2040 than the optimal tax would be if implemented in 2000. This tax rate will curb production, so it does not need to increase more after that. The deterministic steady state in this case is an aggregated level of about 400 p.p.m. with a production equal to 1.75 p.p.m.

In the "worst case" scenario where the damage already today represents as much as 4.6 per cent of GWP, production must be curbed when the aggregated carbon level reaches 420 p.p.m. If implementation is postponed, the optimal rate must be 225 per cent higher already in 2020, or it must increase by as much as 11 per cent a year. A rate 225 per cent higher than the optimal rate in 2000 will curb production and therefore does not need to increase any more. The reason why this is lower than in the previous case, is that the optimal rate in 2000 is higher when the damage is higher. The deterministic steady state is now well below 400 p.p.m. and emissions must be reduced to 1.5 p.p.m. per year in the optimal long run. In this worst case scenario there is a strong indication that either the tax will not be introduced at all or an insufficient tax will be introduced. That is, either too little too late or, at worst, nothing at all.

As the basic assumptions about supply- and demand elasticities are somewhat uncertain, sensitivity analysis with respect to the parameters in the supply and demand functions was performed. Supply elasticity is assumed to be more uncertain than demand elasticity and is therefore varied to +/- 0.55 whereas demand elasticity is varied by +/- 0.1. In this sensitivity analysis the assumption about damage corresponds to the intermediate case above, namely that today's damage is equivalent to 2.8 per cent of GWP. As expected, it is found that optimality regarding the physical variables, that is emissions and aggregated carbon concentration, remain unchanged when the elasticities change. It is only the tax rate that changes. The expected qualitative effect of changing the elasticities is that when supply and demand becomes less elastic this means that producers and consumers are less responsive to changes in price. The tax in this model is a corrective tax designed for the purpose of internalizing an externality where the objective is to achieve a certain goal given in physical units. When the agents in the economy become less responsive to changes in price, a higher price (or tax) is needed to achieve the goal. The results from the sensitivity analysis confirm this expectation, and it is also seen that tax rates are quite sensitive to changes in the elasticities. Only by changing the demand elasticity by +/-

0.1, the tax needed to curb all emission increases from 800 to almost 1000 with reduced elasticity and to less than 700 when the elasticity increases. In other words, the relative change in the tax rate is 15 - 25 percent. When the supply elasticity is changed by +/- 0.55 the effect is even more pronounced. Then the maximum tax rate goes up to 3750 per cent (when the supply elasticity is 1.1) and down to 225 per cent (when the supply elasticity is 2.2).

All in all, the main message in this report is that it may possibly be very expensive to postpone implementation of a carbon tax as the tax rate may have to increase by up to 30 per cent and more for each year implementation is postponed in order to recover optimality in the most pessimistic cases. In the more optimistic cases an increase of down to 0.5 per cent per year may be sufficient. Another important message is that more empirical research on the supply- and demand conditions are necessary along with more research on the actual magnitude of the damage caused by climate change.

References

[1] Boden, Marland (2008), see http://cdiac.ornl.gov/ftp/ndp030/global.1751_2007.ems.

[2] Burgers, van Leeuwen, Evensen (1998). Analysis scheme in the ensemble Kalman filter. Monthly Weather Review 126(6): 1719 – 1724.

[3] Evensen (1994). Sequential data assimilation with a nonlinear quasi-geostropic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research 99(C5): 10143 – 10162.

[4] Evensen (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean Dynamics 53: 343 – 367.

[5] Evensen (2009). Data Assimilation: The Ensemble Kalman Filter. Second edition. Springer-Verlag, Berlin Heidelberg.

[6] Etheridge, Steele, Langenfelds, Francey, Barnola, Morgan (1996). Natural and anthropogenic changes in atmospheric CO_2 over the last 1000 years from air in Antartic ice and firn. Journal of Geophysical Research 101: 4115 – 4128.

[7] Etheridge, Steele, Langenfelds, Francey, Barnola, Morgan (2001). Law Dome Atmospheric CO_2 Data, IGBP PAGES/World Data Center for Paleoclimatology, Data Contribution Series #2001-083, NOAA/NGDC Paleoclimatology Program, Boulder CO, USA, see ftp://ftp.ncdc.noaa.gov/pub/data/paleo/icecore/antarctica/law/readme_law_CO_2.txt

[8] Tans, Keeling (2011), see http://www.esrl.noaa.gov/gmd/ccgg/trends/.