March 24, 2005

Population, uncertainty, and learning in climate change decision analysis

Paper presented at the Annual Meeting of the Population Association of A merica,
Philadelphia, USA
March 31-April 2, 2005

Brian O'Neill (1,2) and Warren Sanderson (3,1)

1 International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria
2 Watson Institute for International Studies, Brown University, Providence, RI, USA
3 State University of New York, Stony Brook, NY 11794-4384, USA.

Abstract

The question of whether to act now or wait to learn more is central to the climate change issue. Previous work has reached no firm conclusions on either the direction or the magnitude of the effect on optimal emissions reductions of incorporating the potential for learning in climate change decision analysis. Here we use a simple model of future emissions and costs of reductions to investigate how learning about the outlook for future population growth could affect optimal climate policy. We draw on recent work showing that, because population growth is path dependent, we can learn about the long term outlook for population size by waiting to observe how population changes in the short term. We find that learning about population growth can affect optimal policy, and that the timing and scope for learning are a key determinant of the magnitude of the effect.

1. Background

The question of whether to act now or wait to learn more is central to the debate over climate change policy. The climate change issue is characterized by both long timescales – today's emissions of greenhouse gases will affect climate for decades to centuries – and substantial uncertainties in climate impacts on society and costs of emissions reductions. Many argue that it would be beneficial to wait to learn more (and reduce uncertainties) before deciding whether, and how much, to reduce emissions. This strategy would avoid investments in emissions reductions that may turn out to be unnecessary. Others argue that reductions should begin now, because if climate change turns out to be serious, we would later regret not acting early. The question of how the potential for learning about various aspects of the problem affects today's optimal decision remains unresolved (Webster 2002).

It is possible that learning about the outlook for future population growth could impact optimal near-term climate change policy decisions. Population uncertainty matters because if population is higher, then all else equal, emissions will be higher and meeting a given emissions reduction target will require larger reductions. If population is lower, meeting the same target will require smaller reductions, or perhaps even none at all. Therefore it is possible that learning about the outlook for population could substantially affect decisions about the best climate policy to pursue in the short term. For example, if 10 or 20 years from now we will know much better whether to expect lower or higher populations over the rest of the century, we might make a different decision about the most appropriate emissions reductions to make today.

To investigate the effect of population uncertainty and learning on optimal climate change policy, we use a very simple model of emissions reduction costs and of future emissions of carbon dioxide as a function of population size. We also use the probabilistic population projections developed at IIASA to model the potential for learning about population over time. We compare three types of optimal solutions (i.e., solutions that minimize the costs of meeting a specified environmental target): (1) the certainty equivalent case, where the uncertainty distribution in population is replaced by its expected value; (2) the optimal solution under full uncertainty; and (3) the optimal solution under uncertainty with learning. In this third case, we solve a two-period, sequential decision-making problem, in which population uncertainty in the second period is partially resolved, contingent on the population path followed in the first period. The conditional population projections for this component of the analysis are derived from the IIASA population projections.

In section 2, we discuss in more detail what we mean by population learning, and illustrate how much learning can reasonably be expected to take place over the coming decades according to the IIASA projections. In section 3 we describe the model we use for emissions and mitigation, and the process of solving for an optimal emissions reduction policy. Section 4 describes the results of our three types of analyses, and section 5 contains concluding thoughts and directions for future work.

2. Learning about future populations by waiting and observing

One thing we know about the future is that it is deeply uncertain. Population forecasters have had enough experience with phenomena such as unpredicted baby booms and busts, and larger-than-anticipated life expectancy increases in some countries coupled with spectacular unanticipated life expectancy decreases in others, that many of them now appreciate the limitations of their art. Population forecasts are, more and more, being produced in a probabilistic framework. These forecasts allow for the possibility of baby booms and busts and for the possibility that life expectancies will increase rapidly in some places and decrease rapidly in others.

One important advantage of using probabilistic forecasting, besides providing a more realistic assessment of what we can say about the future, is that it allows us to study how much we can learn about long-term demographic outcomes by waiting a while and observing outcomes in the short-run (Sanderson, et al. 2004). Deterministic forecasts cannot teach us about learning. If we believe the deterministic forecast, there is nothing to learn. We know what the future will be and waiting provides us with no additional information. The UN produces something that looks like a rough assessment of uncertainty by providing three deterministic paths called low, medium, and high variants (see United Nations (2005), for example). These provide almost no scope for learning either. All uncertainty is resolved after the first five-year time step.

In probabilistic population forecasts, there is uncertainty at each moment of time and the interaction of that uncertainty with population size and age structure provides the possibility of learning about the long-term by observing short-run realizations. The simplest way to see this is to think of an example that has nothing to do with population but uses the metaphor of coin flips.

Imagine that we flipped a fair coin every year for a century and recorded the cumulative number of *heads*. At the beginning of the century, we know that expected number of *heads* is 50, the standard deviation of the number of heads is 5.0 and a 95 percent prediction interval between 40 and 60. Imagine now that we waited for twenty years and we observed only seven *heads* during that interval. Our prediction of the expected cumulative number of *heads* occurring by the end by the end of the century would be 47 (the 7 that have already occurred, plus 40 that would be expected to occur over the remaining 80 years), with a standard deviation of 4.5 and a 95 percent prediction interval between 38 and 56. By waiting twenty years and observing what transpired, we learned that the situation at the end of the century was likely to be different from what we expected at the beginning of the century.

For some purposes, the chances of outcomes in the tails of the distribution are important, and in this example we learn about these outcomes as well. For example, at the beginning of the century, we would say that the probability of observing 56 or more *heads* within 100 years was around 10 percent. Twenty years later, conditional on observing only seven *heads*, we would assess that probability at only around 3 percent.

For some outcomes, however, waiting and observing will not teach us anything. For example, at the beginning of the century, we would predict that the probability of observing 4, 5, or 6 *heads* during the decade 2090-2099 would be around 0.66. Imagine that we waited for twenty years and observed four *heads* in the first decade and three in the next. If we assess the probability of observing 4, 5, or 6 *heads* during the decade 2090-2099 from the later vantage, we would find it to be around 0.66 again. We would have learned nothing about that probability by waiting and observing.

How much we can learn from waiting and observing depends on extent of autocorrelation in the series of interest. The series of the cumulative number of *heads* is highly autocorrelated. The probabilities of observing between 4 and 6 *heads* in a single decade are not. The extent of autocorrelation depends upon both the nature of the random inputs (in the example above these are the coin flips) and the structure in which those random inputs play a part (in the example these are either the cumulative number of *heads* or the probability of observing 4 to 6 *heads* in a decade).

In the case of population, the structure is more complex than in the coin flipping example. First, probabilistic forecasts begin with a known distribution of the population by age and gender. This initial distribution plays an important role in the subsequent evolution of population size and structure, but has no counterpart in the coin-flipping example. Second, the random inputs – fertility and mortality rate schedules in the case of population – are themselves autocorrelated. This is how we account for baby booms and busts and for extended periods of reduced life expectancy due to HIV/AIDS, for example. In contrast, the outcomes of coin flips were assumed to independent of one another. Third, fertility and mortality rates do not affect population size directly, the way an observation of a *head* adds to the cumulative number of heads, but indirectly through their interactions with the age structure. Finally, in dealing with world population, we have to consider temporally correlated region-specific heterogeneity in fertility and mortality patterns. In the coin flipping example, no forms of heterogeneity were included.

The scope for learning about the distribution of future population sizes is substantial for three reasons. First, unanticipated changes such as baby booms and busts, and the spread of HIV/AIDS tend to remain for a while. For example, in Sub-Saharan Africa, reduced life expectancies because of the HIV/AIDS pandemic now seem likely to persist for a quarter of a century or more. Our knowledge of the current situation with respect to HIV/AIDS provides us with information about the likelihood of low life expectancies in Sub-Saharan countries for some years to come. Autocorrelation in the fertility and mortality inputs enhances the possibility for learning.

Second, persistent changes in fertility and mortality lead to changes in age structures that have long-run effects on population size. One well-known example of this is the baby boom in the United States. When the baby boom generation was in their main childbearing years, it had echo effects on the number of births. Demographers use the word "momentum" to describe the phenomenon of population growth due to deviations of a population's age structure from its long-run stable pattern. The existence of "momentum" effects is an important reason that we can learn about long-run population sizes from observing short-run movements in fertility and mortality.

Finally, there is a long-run size effect. When they have identical initial age structures, a larger population, experiencing the same fertility and mortality rates as a smaller population, will always remain larger. When momentum effects cause a population's size to increase or decease, those changes tend to be persistent, even when their initial causes have long vanished.

In studying the implications for greenhouse gas emissions from population learning, we use the probabilistic population forecasts from Lutz, Sanderson, and Scherbov (2001) (see also Lutz, Sanderson, and Scherbov (2004)). Table 1 presents the most relevant features of those forecasts.

Panel A shows the forecasted distribution of world population sizes for the years 2000, 2010, 2020, and subsequently at twenty year intervals to the end of the century, assuming the initial age and sex distribution observed in 2000. The world's population was 6.055 billion people in 2000. The median forecasted population is 6.828 billion in 2010, increases to a plateau of around 8.9 billion in 2060-2080 and then falls to 8.4 billion by

2100. The forecast shows that in 2060, for example, there would be a 1.55 percent chance that the world's population size would be below 6 billion, a 25.75 percent chance that it would be between 8 and 9 billion and a 3.20 percent chance that it would be above 12 billion.

In Panel B, we perform the thought experiment of waiting to 2010, observing whether or not the population was above or below the median forecasted in 2000 (6.828 billion), and then viewing the continuation paths of population, conditional on which half of the distribution we observed. The rows marked "Lower" refer to all population paths that were in the lower half of the population distribution in 2010, while the rows marked "Upper" refer to those that were in the upper half.

Panel A shows that the median forecast world population is 8.4 billion in 2100. Panel B shows that ten years later, population paths that were below the median (in 2010) result in a median population size of 7.7 billion by 2100, while those that were above the median (in 2010) evolve into a distribution with a median of 9.3 billion in 2100.

Someone who concentrates only on the median might think that the change in the median forecast for 2100 from 2000 to 2010 indicates either a mistake or a substantial change of assumptions. After all, how else would it be possible to make a median forecast in 2000 of the world's population in 2100 of 8.4 billion and then, only ten years later, revise the forecast to either 7.7 billion of 9.3 billion? But, in fact, there is nothing wrong nor are there any changes of assumptions in this thought experiment.

It is incorrect to think about a probabilistic population forecast only in terms of its median. The forecast of the world's population in 2100 made in 2000 shows a 42.8 percent chance that the population would be 8 billion or below and a 33.9 percent chance that it would be 9 billion or above. Clearly a population of 7.7 billion or 9.3 billion in 2100 could not be considered to be unlikely. The large changes from a forecasted median of 8.4 billion in 2100 to either a forecasted median of 7.7 billion or 9.3 billion, rather than indicating a problem, show us the power of learning.

If we wait until 2010 and forecast again to 2100, we would obtain quite different distributions of outcomes. If we were in the lower half of the forecasted distribution in 2010, we would assess the chances of ending the century with a population of 8 billion or below at 56.3 percent. If we were in the upper half of the distribution in 2010, we would estimate the chances at only 26.1 percent. If we were in the upper half of the forecasted distribution in 2010, then our forecasts would show only a 20.2 percent chance of having a population of 8 billion or less in 2100, while there would be a 55.4 percent chance of a population size above 9 billion.

Waiting for only ten years and observing the outcomes in that decade would have a clear effect on our forecasts for 2100, with potential implications for appropriate greenhouse gas emissions policies.

3. Model description

We use an extremely simple model for our analysis, which takes a reference scenario/cost function approach to future mitigation opportunities and assumes a single homogenous global region. That is, future emissions in the absence of climate change policy are specified exogenously based on a pre-existing scenario for the period 2000-2100. The costs of reducing emissions are modeled on the basis of a cost function which translates reductions in emissions into associated reductions in GDP. "Controlled" emissions, i.e. emissions after mitigation, are then converted to atmospheric concentrations using a simple carbon cycle model. The optimization problem is defined as finding the path of controlled emissions over the course of the century that minimizes the net present value of mitigation costs, subject to the constraint that atmospheric concentrations cannot exceed a given target level. We recognize that such a model is only a caricature of reality, but it serves the purpose of illuminating the key issues and providing insight into the effects of population assumptions on outcomes. We discuss each element of the model in more detail.

Reference scenario

For a reference emissions scenario, we draw mainly on the IPCC IS92a scenario (Legget et al., 1992). This scenario is representative of many mid-range scenarios in the literature, and therefore serves as a useful starting point for analysis. Although the scenario was produced with a multi-region, multi-sector simulation model of energy and land use, the end result can be expressed in the form of the well know Kaya identity (or, in the population literature, as an I=PAT type identity):

$$E = P \times GPD/P \times E/GDP$$

where E is total carbon emissions from fossil fuel use, P is population, GDP/P is per capita GDP, and E/GDP is the carbon intensity of economic production. We use the per capita GDP and carbon intensity outcomes at the global level from the IS92a scenario, but replace the population projection used in IS92a with a recent projection from IIASA (Lutz et al., 2001), since population is our key variable of interest and we wish to draw on the IIASA projections for our analysis. In the IS92a scenario, per capita GDP increases by X% over the century, rising from XXX per capita in 2000 to XXX in 2100. Carbon intensity continues its historical decline, falling at an average rate of Y%/yr over the century. Declines in carbon intensity can reflect several different processes, including increases in energy efficiency, structural change in the economy toward less energy intensive sectors, and fuel switching away from the most carbon intensive fuels such as coal.

Cost function

We adopt a cost function from the Dynamic Integrated Climate-Economy (DICE) model (Nordhaus, 1994), a well known global integrated assessment model. This function expresses costs as a function of fractional reductions in emissions relative to their level in the reference scenario:

$$TC(t) = b_1 \mu(t)^{b_2} Q(t)$$

where TC(t) is total costs in dollars, μ(t)is the control rate (fractional reduction in emissions), Q(t) is total GDP in dollars, and b_1 and b_2 are parameters. The functional form and parameter values for this cost function were specified in Nordhaus (1994) in order to represent the results of a set of mitigation cost studies taken from the literature. We adopt the same parameter values, $b_1 = 0.0686$ and $b_2 = 2$; i.e., costs rise with the square of the control rate, and an emissions reduction of 10% relative to the reference scenario costs 0.07% of GDP. Costs are measured in terms of deadweight loss to the economy, not in terms of financial costs of the mitigation measures themselves.

Carbon cycle model

Integrated assessment models used to perform uncertainty analysis typically use simple, reduced form models of the global carbon cycle in order to reduce computation time. However in some cases these simplifications have reduced the accuracy of the models (in terms of their ability to reproduce the results of more sophisticated models) so much that the results of the integrated analysis were substantially affected under some conditions (particularly when discount rates are assumed to be low or zero; see e.g. Joos et al., 1999; Kaufmann, 1997; Shultz and Kasting, 1997). We adopt the reduced form model developed by Shultz and Kasting (1997; hereafter SK), which was specifically developed to improve on carbon cycle components of earlier IA models. The model is of the form

$$dM(t)/dt = \beta\, E(t) - M(t)/\tau$$

where M(t) is the atmospheric CO_2 content in excess of the pre-industrial (steady state) value, E(t) is anthropogenic emissions, β is a parameter intended to capture rapid atmospheric removal processes, and τ is a decay constant intended to capture relatively slow atmospheric removal processes. This one-equation model was developed by Nordhaus (1994) for use in his widely-used IA model, DICE, and was calibrated to historical emissions and concentration data. SK pointed out that, based on current understanding of the carbon cycle, its behavior in the future would likely be substantially different than it has been in the past. They proposed that these differences be captured by making the two parameters in eq. (1) functions of both total cumulative emissions and average annual emissions, or

$$\beta(t) = a + b\, E_{tot}(t) + c\, E_{av}(t)$$
$$\tau(t) = d + e\, E_{tot}(t) + E_{av}(t)$$

where $E_{av}(t)$ is the annual average CO_2 emissions, and $E_{tot}(t)$ is the cumulative CO2 emissions. They calibrated this model to future emissions and concentration scenarios generated using their own more complex carbon cycle model, and found it produced substantially different atmospheric CO2 projections than did the original DICE carbon cycle model: the DICE model greatly underestimated future CO2 concentrations over long (>100 year) time horizons.

We adopt the SK model here. We have re-calibrated the model parameters (a, b, c, e; we keep d=1.0 as specified in the original SK model) using historical emissions data, and future emissions and concentration data from carbon cycle model scenarios in which future concentrations are stabilized at levels between 350 and 750 ppm (the "WRE" scenarios; Wigley et al., 1996). Emissions sums and averages in the equations for beta and tau are taken over the period 1750 - t, 1750 being the year anthropogenic emissions are assumed to begin. Historical data on emissions from fossil fuel and from land use change are from Marland et al. (2003).

Objective function

The model is run by solving for the controlled emissions path over time that minimizes mitigation costs. Specifically, the objective function is written as

$$\text{Min} \sum_{t=2000}^{2100} \frac{TC(t)}{P(t)} (1+r)^{2000-t}$$

$$\text{s.t. } C(t) \le \theta$$

where TC(t) is total mitigation costs, P(t) is population, r is the discount rate, C(t) is atmospheric CO_2 concentration, and θ is the stabilization target. Two aspects of this function distinguish it from most others used in IA models of climate change. First, mitigation costs are minimized, rather than maximizing the utility of consumption. We do not employ a utility function, in order to keep the model as simple as possible. Second, it is per capita costs, rather than total costs, that are minimized. We take this approach because unlike the large majority of other analyses, we will be optimizing over states of the world (SOWs) with different population pathways. In such cases, using total costs makes little sense, because SOWs with higher population growth will have larger costs simply due to the larger scale of the economy. Thus an optimization based on minimizing total costs will seek solutions which favor emissions control in SOWs with large populations more so than in those with smaller populations.

Note also that we restrict our period of analysis to 2000-2100, since our population and emissions scenarios are defined only over that period. The summation of costs in the objective function is carried out using time steps of 10 years, and all variables are assumed to remain constant over each time interval. For variables defined from pre-existing scenarios (IS92, IIASA), the value at the beginning of the decade is used as the constant annual value over the whole decade. This is a very rough discretization of the problem, but is useful in that it reduces the number of variables over which to optimize (i.e., decadal control rates), and it makes use of the 2100 information for the population projections without requiring any more sophisticated numerical scheme than the very simple one used here. Thus the costs are calculated over 10 decades (2000-2110), and concentration values are calculated to 2110. Emissions reductions are restricted to be zero in the first decade, so that total emissions are consistent with recent observations.

Optimal emissions vs. optimal taxes

In economic models applied to the climate change issue, there are two basic approaches to calculating optimal mitigation policies: finding optimal pathways for carbon tax rates over time, or finding an optimal pathway for emissions over time. An optimal emissions pathway can be thought of as representing a policy based on mandatory emissions caps, perhaps with an emissions trading scheme that would result, under ideal conditions, in equal marginal reduction costs across regions. In the deterministic case -- i.e., a single given scenario without uncertainty -- the two approaches are equivalent: an optimal carbon tax pathway can be translated into the optimal emissions pathway, and vice-versa. However, under uncertainty, the two approaches diverge.

Consider our application to uncertainty in future reference emissions driven by uncertainty in population. An optimal carbon tax pathway will produce a range of controlled emissions pathways: we know what tax rate will be imposed over time, but we do not know what reference emissions it will be applied to, and therefore we do not know what controlled emissions path will result. In contrast, an optimal emissions path will produce a range of carbon tax pathways: we know what controlled emissions will be (assuming compliance), but we do not know what tax rate would need to be applied to achieve the desired level of controlled emissions, because we do not know what reference emissions will be. Within a stabilization framework, the approach of solving for optimal taxes would imply uncertainty in whether the stabilization target will actually be met. The problem formulation would therefore require specifying a chance constraint, i.e. a maximum probability of exceeding the stabilization level. In the case of an optimal emissions path, the stabilization constraint can be specified with certainty. It is worth noting, however, that it is possible that this approach can produce outcomes in which concentrations actually fall below the target. It is possible that some part of the uncertain range of reference scenarios will lead to concentrations below the stabilization level. In those SOWs, the optimal controlled emissions path will not be followed since it will lie above the reference emissions pathway.

In this analysis we solve for optimal emissions paths, saving an analysis of optimal tax rates for future work. In either a deterministic case or under total uncertainty (i.e., no learning), a one-shot decision must be made at the beginning of the period on what time path of emissions caps to specify for the century. Costs are calculated for each possible reference emissions pathway, and the expected value of these costs is minimized, assuming each reference emissions path is equally likely. If, in any time step, a particular reference scenario falls below the optimal path, costs are assumed to be zero since no reductions are required. Whenever the reference path rises above the optimal path, costs are incurred to reduce emissions to the level of the optimal path in that time step. This occurs even if the reference emissions path rises above the optimal path only temporarily, and the reference path as a whole would not lead to concentrations that exceed the stabilization target. The rationale is that in the total uncertainty case we do not ever know what reference emissions path we are on, and the decision on emissions caps is made at the start of the period and then followed throughout the period. Thus there can be no taking into account of the special circumstances of any given path, because we do not know them.

4. Results

We discuss the results of three analyses: deterministic optimal solutions, optimal solutions under uncertainty in population growth, and optimal solutions under uncertainty with learning by observation.

Deterministic optimal emissions: a sensitivity analysis

As a starting point, we solve for the optimal emissions path given the IS92a scenario, with the median IIASA population projection, as a reference pathway. In addition, we carry out a sensitivity analysis of this result in order to illustrate how population and the other main variables contribute to it. Finally, we show the effect on deterministic results of varying population growth across a plausible range, and also of different stabilization levels. This analysis illustrates the effect of population growth on optimal emissions, which can serve as a basis for understanding the results of our analysis with uncertainty, and then with learning, described later.

Figure 1 shows the optimal emissions path for the IS92a scenario assuming stabilization of atmospheric concentrations at 500 ppm (parts per million), along with four optimal solutions with identical assumptions except in each case one factor in the reference scenario is held constant. The optimal result for the IS92a scenario shows that emissions occur mainly in the beginning of the century, with a peak in about 2030 and a subsequent decline to well below current levels by the end of the century. This implies that while some emissions reductions are made early in the century, most occur many decades into the future. The additional four cases demonstrate that this result is most sensitive to the discount rate. Using a zero discount rate instead of a rate of 3%/yr drastically changes the optimal path from one in which emissions are greatest early in the century to one in which emissions are lowest early and rise over time. This occurs because discounting favors delayed emissions reductions, when they are cheaper in terms of net present value of costs -- a well known feature of least cost mitigation paths.

The second largest sensitivity is to population. If population growth is assumed to be zero rather than following the median IIASA projection, optimal emissions are substantially lower over the first half of the century (and higher over the second half). This occurs because of a cost dilution effect: with a growing population, it is preferable to delay emissions reductions to later when there is a larger population to pay for them (recall that is it per capita costs that are minimized in our formulation of the problem). Thus when there is no population growth, the incentive to delay is removed and optimal emissions are lower early in the century. The third largest sensitivity is to the carbon cycle. With the SK model, later reductions are favored because they have a larger impact on concentrations at the end of the century (the "carbon bonus" (Wigley et al., 1996)). If the carbon cycle model is modified to remove this effect so that reductions at any point in time have the same effect on concentrations, then the carbon bonus to early emissions is lost, and the optimal solution is to make greater reductions early, and smaller reductions late in the century. Finally, the figure shows that there is little sensitivity to carbon intensity. If carbon intensity in the reference scenario is assumed to stay constant rather than falling substantially over time as in IS92a, there is essentially no effect on optimal emissions.[1]

To further explore the effect of alternative population growth paths, we calculate optimal emissions for high, medium, and low population paths (taken to be the 90th, 50th, and 10th percentiles of the IIASA population distribution, respectively), with all other variables following their reference paths. Figure 2 shows that the higher the population growth path, the more emissions reductions are delayed to later in the century. The most important reason for this effect is, again, cost dilution: reductions are cheaper on a per capita basis whenever population is larger. In the high population growth scenario, the size of the population grows the most over time, and the advantages of later reductions are the largest. In contrast, in the low population growth scenario, it pays to make more reductions now (relative to medium growth case) because the advantage of delaying reductions to a time when population is larger is diminished.

The figure also shows that the difference in optimal emissions across the three different population paths considered is noticeable although not extraordinarily large. For example, in 2030, optimal emissions range from about 9.5 - 10.5 GtC/yr depending on the population path. In each of these deterministic cases, the optimal emissions level early in the century is chosen knowing already what population will be in the long term. When we move to the case of uncertainty and of learning, the problem will be to choose optimal emissions before knowing what long term population size will turn out to be. Figure 2 is an indication of the range of possible choices; an optimal choice under uncertainty is likely to lie within the range of deterministic results.

The effect of population on optimal emissions is sensitive to the stabilization level. Stabilization at 500 ppm is a relatively low target, and offers little flexibility in emissions paths: substantial reductions early in the century are necessary no matter what, otherwise reductions later in the century will need to be extremely large (and expensive) in order to meet the concentration target. With a higher stabilization target, the scale of cumulative reductions is smaller, and there is more flexibility in when reductions can be made.

To illustrate, Figure 3 compares optimal emissions assuming either low or high population growth for stabilization at 500 to the same cases for stabilization at 650 ppm. In the 650 case, the reference scenario assuming low population does not require any emissions reductions at all. Although the high emissions baseline does require reductions, controlled emissions are still higher than in the low emissions case until the end of the century. The uncertainty in optimal emissions in the medium term (30-50 years) due to population growth is higher than in the 500 ppm stabilization case: in 2030, the range of

[1] These sensitivies are dependent on the assumed baseline scenario, here taken to be IS92a. Relative to different baselines, the order of importance of the various factors can change.

optimal emissions is about 11-12.5 GtC/yr across the three paths, and expands to 11-13.5 GtC/yr by 2050.

An important alternative way of expressing these results is in terms of costs required to meet these emissions pathways. Figure 4 shows marginal costs of emissions reductions over time for the same four scenarios shown in Figure 3. Comparing the difference between the pairs of results for each stabilization level to the range of results represented by each pair indicates the relative importance of population uncertainty and differences in stabilization level. The effects of these two uncertainties are of comparable size, with the stabilization level somewhat more important early on, and the population path more important later. The marginal cost indicates the carbon tax that would need to be imposed on the economy in order to produce the optimal emissions paths shown in Figure 4. For stabilization at 650, this tax rate could be zero over the entire century if population turns out to be low (and no reductions are required). If population is higher, the tax rate should rise to $50/tC by 2050. For stabilization at 500 ppm, the tax rate would need to be higher in all cases, ranging from about $90 - $140/tC in 2030 depending on population growth, and $210 - $360/tC by 2050. Thus uncertainty in population growth translates into an uncertainty of about $50/tC in 2030 and $100-$150/tC by 2050 -- a substantial effect.

Optimal emissions under population uncertainty

Next, we solve for the single optimal emissions path given uncertainty in future population growth. To find the optimal solution, we minimize the expected value of the sum of discounted mitigation costs per capita. The expectation is calculated over a random sample of 100 possible future population growth paths drawn from the full distribution of IIASA results. Given the population projection methodology, each of these possible population futures is equally likely.

Figure 5 shows that optimal emissions under uncertainty are slightly higher than in the deterministic case assuming medium population growth, with the difference between the two solutions peaking at 0.4 GtC/yr in 2040. The reason for this is that slightly higher emissions hedges against the possibility of high mitigation costs necessary if population growth turns out to be high. The cost consequences of population turning out to be higher or lower than the expected value are asymmetric, partly because the cost function is nonlinear (increasing marginal costs with higher population, and therefore larger emissions reductions) and partly because it is bounded from below by zero. Once population falls low enough to not require any emissions reductions, even lower population has no additional consequences for costs.

Uncertainty in costs is very substantial. Figure 6 shows the 50th and 90th percentiles of the distribution of marginal costs associated with the single optimal emissions path (the 10th percentile of this distribution is at zero). Note that in half the cases (i.e. there is a 50% chance that) tax rates can stay below $10/ton over the first 50 years if this particular pathway of optimal emissions caps is adopted. However there is a 10% chance that they would rise to over $140/tC (the dashed line), which would occur if population turns out to be high and the caps are to be met. The optimal emissions path essentially hedges against this possibly very high tax rate being required, while accounting for the substantial chance that a very low or zero rate may be possible.

Optimal emissions under population uncertainty with learning

Our final analysis accounts for the possibility of learning about future population growth by observation. We define a two period decision problem: In the first period, there is complete uncertainty about future population growth, described by the full distribution of outcomes in the IIASA projections. Between period 1 and 2, we assume that we can observe whether the global population size falls within one of three equally probable quantiles; i.e., whether population has turned out be in the upper, middle, or lower third of the projected outcomes at that point in time. In period 2, the uncertainty in the population outlook for the rest of the century is contingent on the observation. For example, if population is observed to fall within the upper third of outcomes, then we assume that population in period 2 will fall within the conditional distribution consisting of the subset of outcomes in the IIASA projections that are in the top third of population size at the time the observation is made.

Figure 7 shows results for two experiments in which learning occurs either in 2020 or in 2030. In both cases, learning has a clear effect on optimal emissions after it occurs. Three diverging optimal emissions paths are shown after the learning point, each contingent on having observed that global population size was in one of the three equally probably quantiles of the population size distribution at the time of observation. The effect of learning on optimal emissions before the time of observation is smaller. Relative to the solution with total uncertainty, the solutions with learning suggest that the anticipation of learning about population growth should lead to slightly higher emissions in the short term; i.e., some emissions reductions should be delayed until later in the century, relative to the optimal solution when there is uncertainty but no learning. For example, optimal emissions with learning are about 0.3 GtC/yr higher than in the total uncertainty case in the year 2020.

To understand this result, recall that in the deterministic cases we found that higher population growth led to generally higher marginal reduction costs over time, and higher short term emissions in the optimal solution, because cumulative per capita costs could be reduced by postponing reductions to later in the century when costs could be shared across a larger population. In the total uncertainty case, we found that the optimal solution was to hedge against the possibility that population would turn out to be high (and therefore that marginal costs would also be high) by postponing some reductions to later, relative to a deterministic case with population at is median value. Now, we find that if we add the anticipation of learning, the optimal near-term solution is to hedge even further against the possibility that population turns out to be high by postponing additional reductions to later in the century. The reason for this can be understood in terms of the relative regret associated with what we learn between periods (Webster, 2002). Inevitably, after learning whether we are likely to be on a higher or lower population growth path for the rest of the century, we will have some regret over our actions in period 1. If we learn that we are on a higher than expected population growth path, we will wish that we had emitted more in period 1 (i.e., postponed more reductions to period 2). It we learn that we are on a lower than expected population growth path, we will wish that we had emitted less in period 1 (i.e., postponed fewer reductions to period 2). It is the relative magnitude of these two regrets that determines the direction of the learning effect on period 1 emissions. Given the assumptions in our model, we would regret more strongly the situation in which we find out that we are on a higher population growth path. There are two reasons for this. First, the mitigation cost function is nonlinear, and marginal costs increase with higher reductions. Therefore, if we find out we are on a higher population path – requiring larger reductions – the increase in costs is larger than the overspending that would have occurred if we find out population will be lower and we require fewer reductions. The second reason is that in a subset of low population outcomes, population is so low that no reductions are required at all, and therefore the (zero) costs are unaffected by the emissions caps. This means that emissions reductions are less sensitive to the emissions caps set over period 1 in low population outcomes, relative to high population outcomes. We would therefore regret less the situation in which we find we are on a lower population path.

Figure 8 shows the implications of learning for marginal costs, for the example in which the observation takes place in 2030. For comparison, it also shows the marginal costs in the analysis of total uncertainty with no learning. Results show that the possibility of learning is valuable mainly because it leads to a reduction in the probability of high mitigation costs. This occurs for two reasons. In the learning case, after 2030 optimal emissions can be adjusted depending on what is learned about population growth. Thus, if population growth turns out to be high, marginal costs will be high, but not as high as they are in the total uncertainty case because the post-2030 mitigation strategy can be adjusted based on the new knowledge. As a result, the 90th percentile of the distribution of possible marginal costs in 2050 is about \$125/tC, rather than \$150/tC in the total uncertainty case. Second, emissions before 2030 are adjusted in the learning case given the anticipation of learning, as described above. Hedging against high population high population outcomes by making fewer reductions in period 1 means that the likelihood of that marginal costs will be high in period 1 is decreased – the 90th percentile of marginal costs is about \$5/tC less in the learning case relative to the total uncertainty case.

5. Discussion

We find that both incorporating population uncertainty, and incorporating changes in that uncertainty over time due to learning by observation about future population growth, leads in the short term to a higher optimal emissions path to stabilization of atmospheric concentrations of CO2. We also find that learning leads to a reduction in the likelihood of facing the highest emissions reduction costs. While the magnitudes of the effects are not extremely large, we consider the results more as a proof of concept based on the simplest possible model. Further investigation is required to explore sensitivity to a number of limitations and simplifications in our analysis. For example, we have assumed a very simple mitigation cost function in which nonlinearities are not strong, and there is little inertia in the energy or economic system. Emissions reductions made today do not have any effect on costs in the future, when in fact maintenance and operation costs could stretch into the future, mitigation costs could have dynamic macro-economic effects, or the experience gained with current reductions could reduce the cost of future reductions. These effects would be important to determining the effect of the potential for learning on emissions reductions. In addition, we have assumed a very specific objective function: minimization of the expected value of per capita reduction costs. If, for example, reducing the risk of high costs, and not just the expected value, is important, learning could have a larger effect since we see already in our results its ability to reduce the risks of high cost outcomes. We also have explored the effect of learning using only a single stabilization level, when in fact a much more accurate framing of the problem would include uncertainty in what stabilization level we will eventually want to reach. We have also used a single world region; a better approach would be to disaggregate by regions, so we can better match which populations we will learn most about with their expected contributions to future emissions. An additional question that remains to be addressed is what the effect of learning would be if climate policy is implemented not as emissions caps, as done here, but in the form of a specified tax rate. As discussed earlier, the two approaches differ under conditions of uncertainty.

Finally, we have also used a fairly rudimentary implementation of population learning, in which our observation can only discern global population size, and only at very poor resolution: falling into one of three broad size categories. In reality, it will be possible to observe much more information than that: global population size can be estimated with much greater accuracy, and information about regional population and fertility rates (another indicator of the potential for future growth) will also be available. More work is required to understand how much, and how fast, we can realistically expect to learn about future population growth.

Nonetheless, the results presented here demonstrate that the potential for learning about our demographic future can affect near term climate change policy decisions, and can also reduce the costs of achieving climate change goals in the long term.

References

Joos, F., Muller-Furstenberger, B., and Stephan, G. 1999. Correcting the carbon cycle representation: How important is it for the economics of climate change? Environmental Modeling and Assessment 4, 133-140.

Kaufmann, 1997;

Leggett, J., Pepper, W. J., and Swart, R. J.: 1992, *Climate Change 1992: The Supplementary Report to the IPCC Scientific Assessment*, Cambridge University Press, Cambridge.

Lutz, W., Sanderson, W., and Scherbov, S., 2001. The end of world population growth. Nature 412, 543-545.

Lutz, W., W. Sanderson, and S. Scherbov, (2004). The End of World Population Growth. In W. Lutz, W.C. Sanderson, and S. Scherbov, ed. *The End of World Population Growth in the 21st Century: New Challenges for Human Capital Formation & Sustainable Development*. London: Earthscan. Pp. 17-83.

Marland, G., T.A. Boden, and R. J. Andres. 2003. Global, Regional, and National Emissions. 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

Nordhaus, W.D., 1994, Managing the Global Commons, MIT Press, Cambridge, MA, USA.

Sanderson, W., Scherbov, S., O'Neill, B.C., and W. Lutz. (2004) Conditional probabilistic population forecasting, *International Statistical Review* 72(2), 157-166.

Shultz, P.A. and Kasting, J.F. 1997. Optimal reductions in CO2 emissions. Energy Policy 25(5), 491-500.

United Nations (2005). World Population Prospects: The 2004 Revision. Highlights. New York: United Nations

Webster, M. 2002. The curious role of "learning" in climate policy: Should we wait for more data? *The Energy Journal.* 23(2). 97-119.

Wigley, T.M.L., Richels, R. and Edmonds, J.A. 1996. Economic and environmental choices in the stabilization of atmospheric CO2 concentrations. Nature 379, 240-243.

Year	World Population Size Intervals (Billions)								
	Below 6	*6-7*	*7-8*	*8-9*	*9-10*	*10-11*	*11-12*	*Above 12*	*Median*
Panel A: 2000 Base									
2000	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	6.055
2010	0.00	85.05	14.95	0.00	0.00	0.00	0.00	0.00	6.828
2020	0.00	8.05	81.45	10.50	0.00	0.00	0.00	0.00	7.538
2040	0.15	3.75	24.85	40.35	25.15	5.20	0.65	0.00	8.525
2060	1.55	7.00	17.45	25.75	22.35	16.25	6.45	3.20	8.935
2080	6.80	10.10	16.00	18.45	18.45	12.55	9.00	8.65	8.890
2100	14.25	14.05	14.45	16.50	12.90	10.45	6.85	10.55	8.413
Panel B: 2010 Base									
	Below 6	*6-7*	*7-8*	*8-9*	*9-10*	*10-11*	*11-12*	*Above 12*	*Median*
2020/Lower	0.00	16.10	83.90	0.00	0.00	0.00	0.00	0.00	7.268
2020/Upper	0.00	0.00	79.00	21.00	0.00	0.00	0.00	0.00	7.787
2040/Lower	0.30	7.50	42.30	42.40	7.40	0.10	0.00	0.00	7.996
2040/Upper	0.00	0.00	7.40	38.10	42.90	10.30	1.30	0.00	9.083
2060/Lower	3.10	12.80	26.90	31.30	17.30	6.80	1.60	0.20	8.256
2060/Upper	0.00	1.20	8.00	20.20	27.40	25.70	11.30	6.20	9.760
2080/Lower	12.50	14.70	22.20	19.80	16.90	7.90	4.10	1.90	8.045
2080/Upper	1.10	5.50	9.80	17.10	20.00	17.20	13.90	15.40	9.816
2100/Lower	22.00	17.70	16.60	17.60	9.90	9.00	3.60	3.60	7.652
2100/Upper	6.50	1.40	12.30	15.40	15.90	11.90	10.10	17.50	9.328

Table 1: Forecasted Distributions of World Population, 2000, 2010, 2020, 2040, 2060, 2080, and 2100.
Panel A uses 2000 as the jump-off year of the forecast. Panel B uses 2010 as the jump-off year of the forecast, conditional on whether a population path was in the Lower or Upper half of distribution in 2010.
Source: Sanderson et al. (2004).

Figure 1: Deterministic optimal emissions for stabilization at 500 ppm assuming the IS92a reference scenario (thick red), and IS92a with constant carbon intensity (thin red), modified carbon cycle (light green); constant population (orange); zero discount rate (blue).

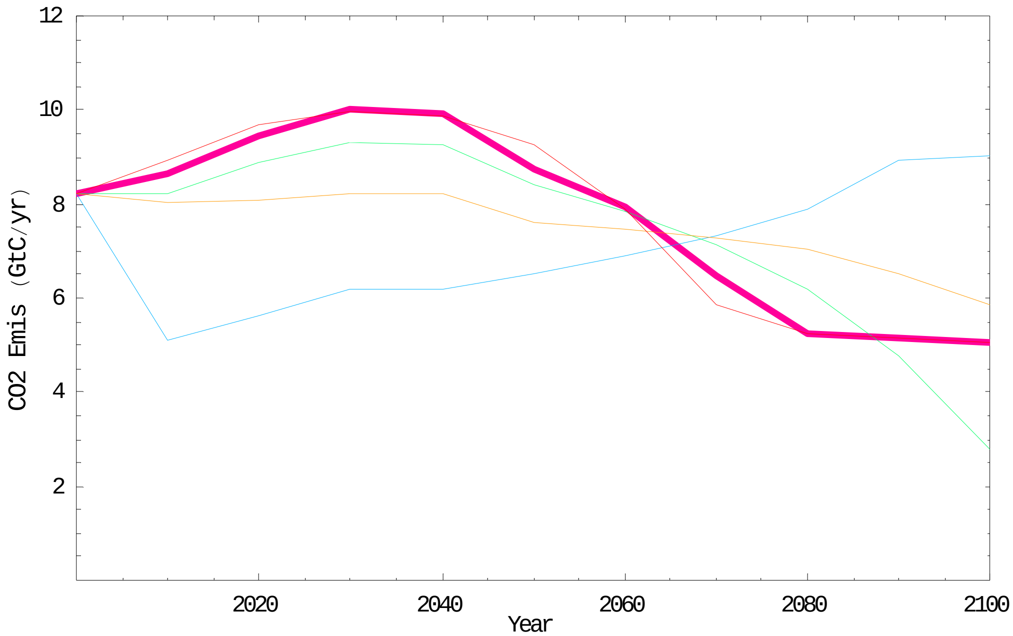

Figure 2: Deterministic optimal emissions for stabilization at 500 ppm, assuming a population growth path that is low (thin red), medium (thick red), or high (thin purple).

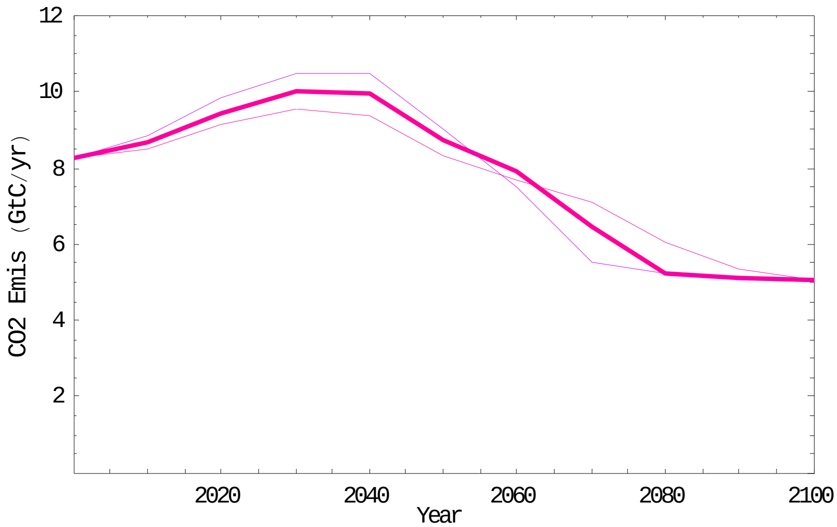

Figure 3: Deterministic optimal emissions paths for alternative stabilization levels and population assumptions. Red: 650 ppm stabilization level; Blue: 500 ppm stabilization level; Dashed: low population; Solid: high population.

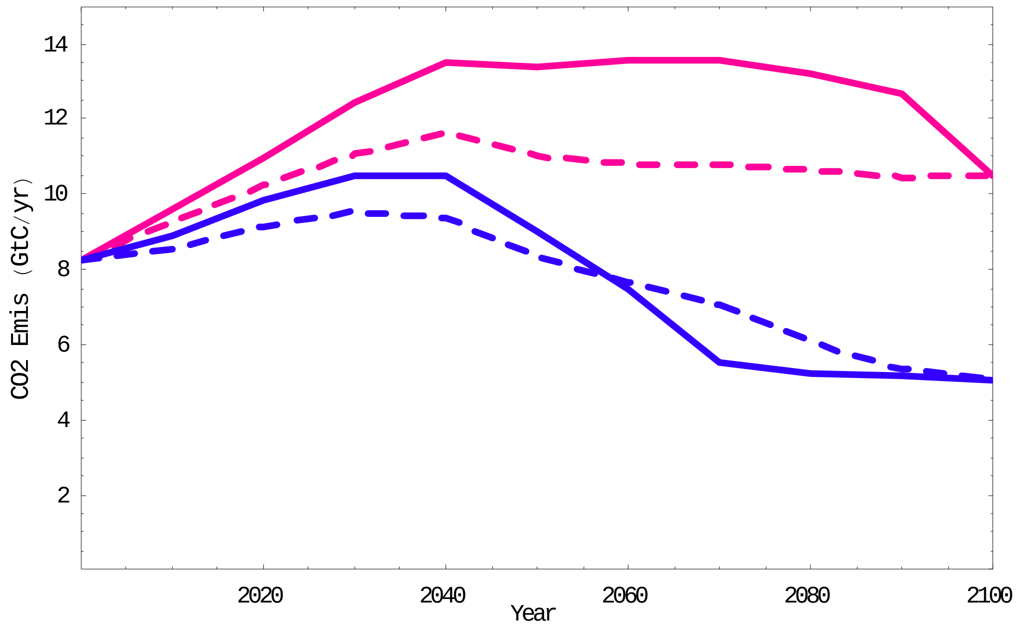

Figure 4: Marginal costs of emissions reductions in dollars per ton of carbon reduced for the four optimal emissions paths shown in Figure 3.

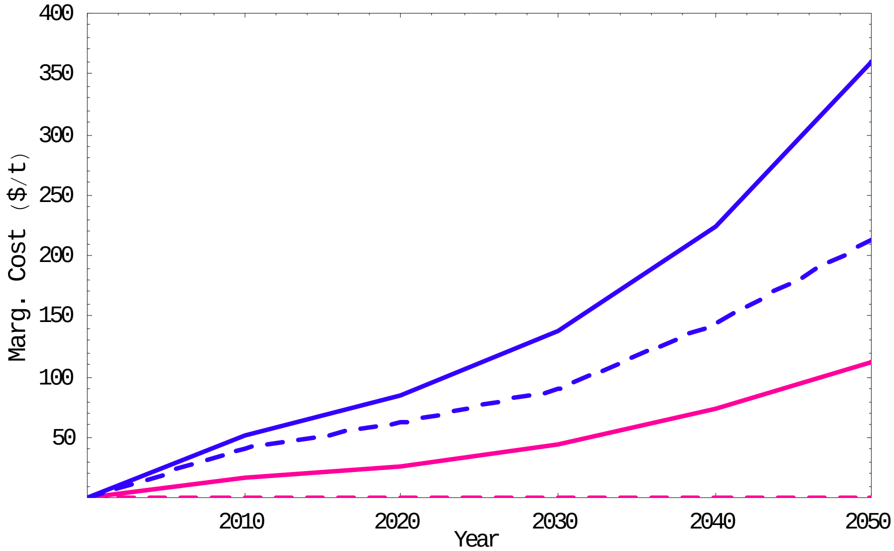

Figure 5: Optimal emissions for stabilization at 650 ppm under uncertainty in population (thick blue line). For comparison, the deterministic results for alternative population assumptions are shown in red: medium population (thick red); low population (thin red); high population (dashed red).

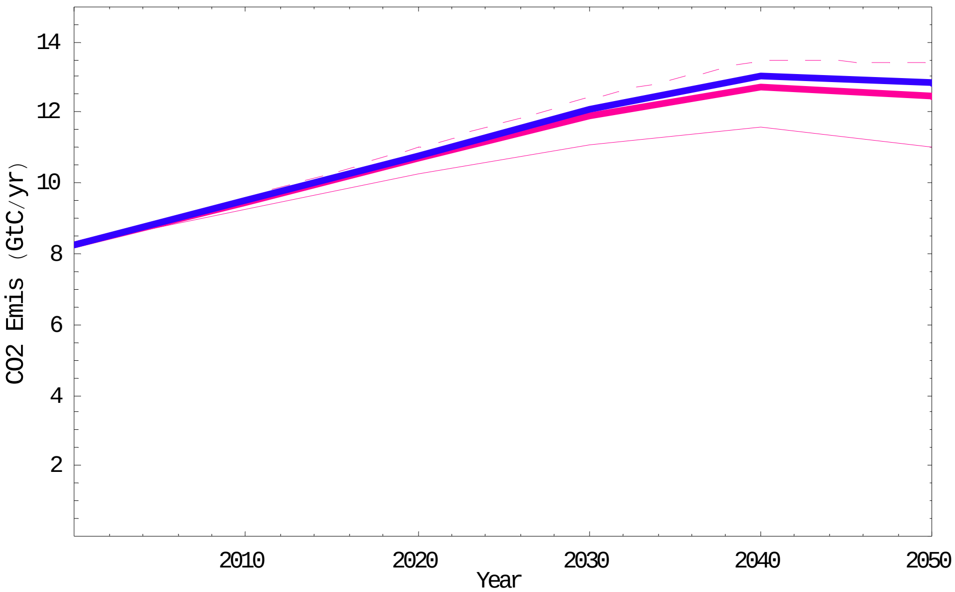

Figure 6: Percentiles of the distribution of marginal reduction costs associated with optimal emissions under population uncertainty for stabilization at 650 ppm. 90th percentile: dashed red; 50th percentile (thick solid red).

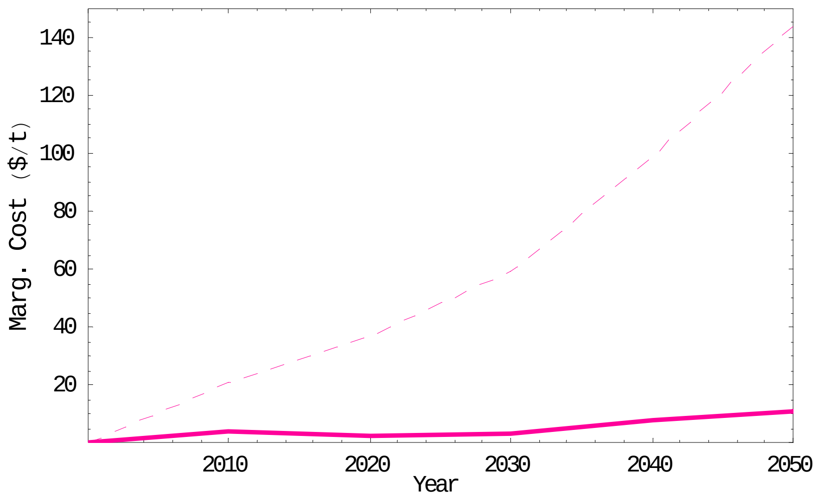

Figure 7: Optimal emissions for stabilization at 650 ppm under uncertainty in population (thick blue line) and assuming learning in 2020 (red lines) or 2030 (green lines).

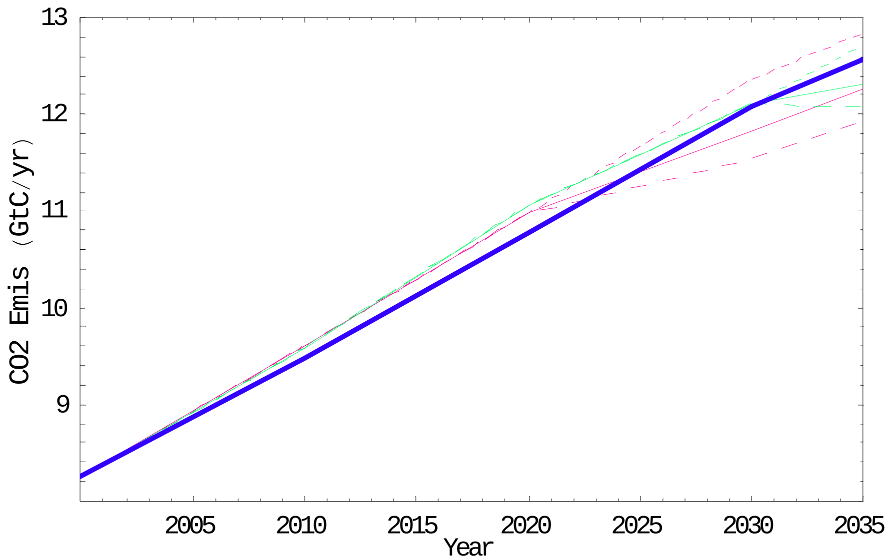

Figure 8: Percentiles of the distribution of marginal reduction costs associated with optimal emissions under population uncertainty (blue lines) and assuming learning takes place in 2030 (red lines) for stabilization at 650 ppm. 90th percentile: dashed; 50th percentile: thick solid.