Steering the Climate System: Using Inertia to Lower the Cost of Policy

By Derek Lemoine and Ivan Rudik*†

Common views hold that the efficient way to limit warming to a chosen level is to price carbon emissions at a rate that increases exponentially. We show that this "Hotelling" tax on carbon emissions is actually inefficient. The least-cost policy path takes advantage of the climate system's inertia to delay reducing emissions and allow greater cumulative emissions. The efficient carbon tax follows an inverse-U-shaped path and grows more slowly than the Hotelling tax. Economic models that assume exponentially increasing carbon taxes are overestimating the cost of limiting warming, overestimating the efficient near-term carbon tax, and overvaluing technologies that mature sooner.

JEL: H23, Q54, Q58

Keywords: climate, hotelling, emissions, inertia, tax, carbon, abatement, dynamics

* Lemoine: Department of Economics, University of Arizona, 1130 E. Helen St, McClelland 401, Tucson, AZ 85721, dlemoine@email.arizona.edu. Rudik: Department of Economics and Center for Agricultural and Rural Development, Iowa State University, 260 Heady Hall, Ames, Iowa 50011, irudik@iastate.edu. This work was supported by the University of Arizona's Renewable Energy Network, with special thanks to Ardeth Barnhart and Stan Reynolds. We thank Haewon McJeon for helpful comments. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

In recent years, several international agreements have committed nations to limiting global warming to 2 degrees Celsius (Jaeger and Jaeger, 2010; Gillis, 2014). The 2015 Paris Agreement even encourages nations to limit warming to 1.5 degrees Celsius. These temperature limits require substantial, costly reductions in carbon dioxide (CO_2) emissions over the next century. Surprisingly, economists have yet to theoretically analyze the emission trajectory that efficiently limits warming to a chosen level. We demonstrate that the efficient policy trajectory postpones emission reductions to take advantage of the climate system's considerable inertia.

If the goal were to limit the accumulation of CO_2 in the atmosphere, then the least-cost policy would price emissions at a level that increases at the the rate of interest plus the rate at which CO_2 "decays" in the atmosphere (Nordhaus, 1980, 1982; Peck and Wan, 1996; Goulder and Mathai, 2000). This least-cost trajectory is commonly called a Hotelling trajectory: if we consider the atmosphere's CO_2-holding capacity as an exhaustible resource whose quantity is fixed by the chosen CO_2 limit, then the least-cost policy depletes the resource (via emissions) according to the analysis of Hotelling (1931). The intuition is as follows. Along a least-cost trajectory, the policymaker must be indifferent to small deviations in the trajectory. Imagine that the policymaker considers deviating by allowing an additional unit of emissions today. Instead of spending money on reducing emissions today, the policymaker would invest those savings and compensate by undertaking additional emission reductions t years in the future. In order to return to the original CO_2 trajectory, the policymaker will not need to reduce future emissions by a full unit because the additional unit of emissions will have decayed at rate δ. By deviating in this fashion, the policymaker has earned interest at rate r over those t years and has also seen the required spending decline at the rate δ of CO_2 decay. In order for the policymaker to be indifferent to this deviation, the marginal cost of emission reductions (i.e., the tax on CO_2 emissions) must grow at rate $r + \delta$.

We show that the marginal cost of emission reductions should follow a qualitatively different trajectory when policymakers aim to limit total warming rather than total CO_2. The reason is that an increase in CO_2 neither immediately nor fully translates into an increase in warming. The climate system displays substantial inertia, warming only slowly in response to additional CO_2.[1] A year's temperature is determined not just by the contemporary quantity of CO_2 in the atmosphere but also by the past trajectory of CO_2. Additional warming incurred by temporarily raising CO_2 cannot be undone simply by returning to the original CO_2 trajectory. By allowing additional warming over the next t years, a policymaker sacrifices some of the braking services provided by the inertia in the climate system. In order to return to the original temperature trajectory, the policymaker must undertake a sufficiently large quantity of emission reductions to bring time t CO_2 some distance *below* its original trajectory. This additional spending offsets the policymaker's earnings from interest and from the natural decay of CO_2. The efficient tax on CO_2 emissions must grow more slowly than exponentially.[2]

[1]For example, interactions with ocean heat sinks mean that the next decades' warming will represent only about 50–60% of the eventual equilibrium warming corresponding to their likely CO_2 concentrations (Solomon et al., 2009). Even if we were to freeze all greenhouse gases at their current concentrations, the climate system's inertia means that we could expect total warming to more than double from the current level (Wetherald, Stouffer and Dixon, 2001).

[2]Policymakers are increasingly discussing "geoengineering" approaches to controlling climate change. These approaches would directly control temperature rather than CO_2, perhaps by shooting reflective particles into the atmosphere. In the appendix, we show that if a policymaker were willing to use this type of technology to achieve a temperature

The presence of inertia in the climate system is valuable for a policy aiming to limit total warming. This value manifests itself in two ways. First, inertia allows the policymaker to delay emission reductions without immediately incurring the full temperature penalty. For any positive consumption discount rate, the temporary disconnect between CO_2 and temperature provides a valuable degree of freedom which the policymaker uses to lower the present cost of policy.[3] Second, inertia allows the policymaker to reduce the cumulative quantity of abatement undertaken over time. By delaying the temperature consequences of additional CO_2, the climate system's inertia allows more time for CO_2 to decay. Even if future abatement costs are not discounted, the policymaker reallocates abatement over time so as to reduce the cumulative quantity of abatement undertaken. In the presence of discounting or of natural decay of CO_2, the climate system's inertia allows for a lower initial tax and reduces the overall cost of the policy program.

Our results highlight a previously unrecognized flaw in estimates of the cost of limiting warming. The primary tools for estimating these costs are multisector market equilibrium models, called "cost-effectiveness integrated assessment models." Some economists criticize this modeling approach for not endogenizing savings or growth. Nonetheless, these models are the preferred tools for estimating the economic implications of proposed policies because they implement detailed representations of energy systems, technologies, and climate dynamics. Other economists criticize this modeling

limit, then its efficient deployment would in fact follow a Hotelling trajectory, with the rate of increase modified by the degree of climatic inertia rather than by the rate at which CO_2 decays.

[3]In particular, we show that the least-cost policy temporarily overshoots the steady-state CO_2 level required by the temperature limit. Wigley (2003), Huntingford and Lowe (2007), and Wigley, Richels and Edmonds (2007) previously suggested that overshoot trajectories might in fact be cheaper ways of achieving climate goals. Subsequent numerical experiments have supported this conjecture.

approach for not optimizing the emission tax by trading off the welfare loss from climate change. Two arguments suggest that cost-effectiveness approaches can nonetheless provide valuable economic analysis. First, we know remarkably little about the harm from climate change. Such ignorance can justify analyzing predefined limits on temperature (Baumol, 1972). Second, global climate agreements are clearly oriented around limiting warming to 2 degrees Celsius. Economic analysis should guide the translation of this goal into policy.

We demonstrate a new first-order problem with cost-effectiveness models' internal logic. These models' detailed structures can prevent them from flexibly searching for the policy trajectory that minimizes the cost of limiting warming. Instead, many of them assume that the cost-minimizing policy trajectory has the modified Hotelling form described above (Bauer et al., 2015). Contrary to common views (e.g., Tol, 2013), we show that this policy path does not minimize the cost of limiting temperature to a chosen level. We show that using the incorrect policy trajectory can lead models to overestimate the cost of meeting a 2°C temperature target by a factor of 10–100. The errors from failing to endogenize savings or from failing to trade off the welfare loss from climate change are unlikely to be as large: in the benchmark cost-benefit integrated assessment model (Nordhaus, 2008), the endogenous savings rate does not vary much across specifications and temperatures below 2°C reduce output by only 1% or less. By implementing policy paths that ignore inertia, computational equilibrium models' results have overstated the minimum cost of achieving temperature limits, overestimated the level of the near-term emission tax consistent with these limits, and overvalued technologies that mature sooner rather than later.

I. Setting

A global planner seeks the least-cost emission path to limit global warming to an exogenous level $\bar{T}$. The setting is in continuous time, with an infinite-horizon planning period. Business-as-usual CO_2 emissions $E > 0$ arise exogenously. The policymaker chooses each instant's quantity of abatement $A(t)$, with the net emissions released to the atmosphere becoming $E - A(t)$. The cost of abatement is $C(A(t))$, where $C(\cdot) : \mathbb{R}_+ \to \mathbb{R}$ is an increasing, twice-differentiable, continuous, and strictly convex function. For ease of exposition, we assume that E and $C(\cdot)$ are stationary, and we ignore potential nonnegativity constraints on abatement and net emissions because they do not bind under the calibrations reported in the main text.

Atmospheric carbon dioxide $M(t)$ is increased by net emissions. CO_2 in excess of the preindustrial concentration M_{pre} decays at rate $\delta \in (0, 1)$:[4]

$$\dot{M}(t) = E - A(t) - \delta \left(M(t) - M_{pre} \right), \tag{1}$$

where dot notation indicates a time derivative. Atmospheric CO_2 generates forcing $F(M(t)) > 0$, with $F'(M(t)) > 0, F''(M(t)) < 0$. Forcing measures the greenhouse effect, which traps outgoing heat. If maintained forever, one unit of forcing would generate $s > 0$ units of warming, where s is a transformation of the parameter commonly known as climate sensitivity. However, climatic inertia means that forcing does not immediately translate

[4]The appendix demonstrates that our primary analytic results are robust to the more complex carbon model of Golosov et al. (2014). The substantive differences resulting from that setting are that cumulative emissions are fixed by the temperature target and that the nonnegativity constraint on net emissions binds in our calibration.

into temperature:

$$\dot{T}(t) = \phi \left[s\, F(M(t)) - T(t) \right]. \tag{2}$$

The parameter $\phi > 0$ controls the degree of inertia in the system. Greater ϕ indicates less inertia. As $\phi \to \infty$, there is no inertia: an instant's forcing completely determines that instant's temperature. As $\phi \to 0$, there is full inertia: temperature never changes, irrespective of forcing. This temperature representation follows Nordhaus (1991) and is a reduced version of the temperature module used in Nordhaus (2008).

The initial time t_0 is given. The initial level of CO_2 is $M_0 > M_{pre}$, and initial temperature is $T_0 < \bar{T}$. Assume that $E > \delta \left(F^{-1}(\bar{T}/s) - M_{pre} \right)$, so that maintaining temperature at $\bar{T}$ requires strictly positive abatement. The policymaker selects an abatement trajectory in order to minimize the present cost of maintaining temperature below the policy target:

$$\min_{A(t)} \int_{t_0}^{\infty} e^{-r(t-t_0)}\, C(A(t))\, dt \tag{3}$$

$$\text{subject to equations (1) and (2), } T(t) \leq \bar{T},\ M(t_0) = M_0,\ T(t_0) = T_0.$$

The policymaker discounts costs at rate $r > 0$. We assume that damages from climate change are negligible for $T(t) \leq \bar{T}$. This approach is consistent with international policy discussions and also with the technology-rich numerical models used to evaluate policy. Including pre-threshold damages would not affect our theoretical insights.

Define $\bar{M}$ as the unique CO_2 concentration compatible with the climate system remaining at $\bar{T}$: $\bar{M} \triangleq F^{-1}(\bar{T}/s)$. The climate dynamics themselves directly imply two important results, proved in the appendix:

PROPOSITION 1:

1) *Along a least-cost path, there exists a time q such that $\dot{M}(t) \leq 0$ for all times $t \geq q$ and $\dot{M}(t) < 0$ for some times $t \geq q$.*

2) *A path constrained by temperature limit $\bar{T}$ can achieve strictly less cost than a path constrained by the corresponding CO_2 limit $\bar{M}$.*

The first result says that a least-cost CO_2 trajectory overshoots the steady-state CO_2 level consistent with the temperature constraint. This occurs because the inertia in the climate system enables CO_2 to temporarily exceed its steady-state level without violating the temperature constraint. Any path that does not take advantage of this ability to overshoot the steady-state CO_2 level cannot be a least-cost path. The proposition's second result follows from the first: because a least-cost path must overshoot its steady-state CO_2 level, indirectly achieving a temperature constraint by directly constraining CO_2 must increase the cost of the efficient policy program.

II. Least-Cost Policy

The policymaker faces a control problem with a pure state constraint. See the appendix for background on such problems. When the state constraint binds, the choice of the control $A(t)$ is completely determined by the constraint. Following Hartl, Sethi and Vickson (1995), write the state constraint as $h^0(M(t), T(t), A(t)) = \bar{T} - T(t) \geq 0$. Totally differentiating with respect to time, we have:

$$h^1(M(t), T(t), A(t)) \triangleq \frac{\mathrm{d}h^0(M(t), T(t), A(t))}{\mathrm{d}t} = -\phi \left[s\, F(M(t)) - T(t) \right],$$

$$h^2(M(t), T(t), A(t)) \triangleq \frac{\mathrm{d}h^1(M(t), T(t), A(t))}{\mathrm{d}t} = -\phi \left[s\, F'(M(t))\, \dot{M}(t) - \dot{T}(t) \right].$$

This state constraint is of order two because the control variable enters at the second derivative with respect to time (via $\dot{M}(t)$). Form the current-value Hamiltonian:

$$\begin{aligned}H(M(t), T(t), A(t), \lambda_M(t), \lambda_T(t)) =& C(A(t)) + \lambda_M(t)\,[E - A(t) - \delta\,(M(t) - M_{pre})] \\ &+ \lambda_T(t)\,\phi\,[s\,F(M(t)) - T(t)].\end{aligned}$$

The current-value Lagrangian is

$$H[t]+\nu(t)\left\{-\phi s\,F'(M(t))\,[E - A(t) - \delta\,(M(t) - M_{pre})]+\phi^2\,[s\,F(M(t)) - T(t)]\right\},$$

where we write $[t]$ in place of the Hamiltonian's full set of arguments. In addition to the transition equations, the initial conditions, and the state constraint, a least-cost trajectory must satisfy the following necessary conditions (Hartl, Sethi and Vickson, 1995).[5] First, it must satisfy the Maximum Principle and the adjoint equations:

$$C'(A(t)) = \lambda_M(t) - \nu(t)\phi\,s\,F'(M(t)), \tag{4}$$

$$\begin{aligned}\dot{\lambda}_M(t) =& (r+\delta)\,\lambda_M(t) - \phi\,s\,F'(M(t))\,\lambda_T(t) \\ &+ \nu(t)\,\phi\,s\left\{F''(M(t))\dot{M}(t) - F'(M(t))\,[\delta+\phi]\right\},\end{aligned} \tag{5}$$

$$\dot{\lambda}_T(t) = (r+\phi)\,\lambda_T(t) + \nu(t)\phi^2, \tag{6}$$

[5]Throughout, we focus on necessary conditions because the well-known scientific finding that $F''(\cdot) < 0$ prevents the application of standard sufficiency conditions. We adapt the necessary conditions from Hartl, Sethi and Vickson (1995) to reflect the minimization objective and to express the multipliers in current-value terms.

where primes indicate derivatives. Second, we have the analogue of the Kuhn-Tucker conditions for the temperature constraint multiplier:

$$\nu(t)[\bar{T} - T(t)] = 0, \quad \nu(t) \leq 0, \quad \dot{\nu}(t) \geq r\nu(t), \quad \ddot{\nu}(t) \leq 2r\dot{v}(t) - r^2\nu(t).$$

Third, we have the jump conditions for the costate variables. Note that once the temperature constraint begins to bind, it must bind forever along a least-cost trajectory. Let τ denote the time at which the temperature constraint first binds, known as the entry time. The appendix shows that a least-cost trajectory cannot approach $\bar{T}$ only asymptotically. Intuitively, a path that maintains $T(t) < \bar{T}$ at all times t is more costly than one that allows slightly more emissions yet still remains weakly below $\bar{T}$. The least-cost path must therefore attain $\bar{T}$ in finite time, which means that τ is finite. The costate variables can jump at τ:

$$\begin{aligned} \lambda_M(\tau^-) =& \lambda_M(\tau^+) - e^{r(\tau - t_0)}\eta_M^2 \phi s F'(M(\tau)), \\ \lambda_T(\tau^-) =& \lambda_T(\tau^+) - e^{r(\tau - t_0)}\eta_T^1 + e^{r(\tau - t_0)}\eta_T^2 \phi, \\ H[\tau^-] =& H[\tau^+], \end{aligned} \tag{7}$$

where $\eta_M^2, \eta_T^1, \eta_T^2 \leq 0$ and where superscript plus and minus indicate right and left limits, respectively. A final set of necessary conditions relates the jump variables η to the constraint multiplier $\nu(t)$:

$$\eta_T^1 \leq -e^{-r(\tau - t_0)}\,\dot{\nu}(\tau^+) + re^{-r(\tau - t_0)}\nu(\tau^+), \quad \eta_M^2 = \eta_T^2 = e^{-r(\tau - t_0)}\,\nu(\tau^+). \tag{8}$$

Now consider least-cost policy once the constraint binds and just before the constraint binds. Because the constraint binds for all $t \geq \tau$, we have

$h^1 = 0$ and $h^2 = 0$ for all $t \geq \tau$. $h^1 = 0$ implies that $M(t) = \bar{M}$, and $h^2 = 0$ implies that $A(t) = E - \delta[\bar{M} - M_{pre})] \triangleq \bar{A}$. From equation (4), we have $C'(A(\tau^-)) = \lambda_M(\tau^-)$ and $C'(A(\tau^+)) = \lambda_M(\tau^+) - \nu(\tau^+)\phi\, s\, F'(M(\tau))$. Equation (7) and the necessary conditions in (8) then imply that abatement is continuous at τ: $A(\tau^-) = A(\tau^+)$. Therefore, as the system approaches time τ, we know that $T(t) \to \bar{T}$, $M(t) \to \bar{M}$, and $A(t) \to \bar{A}$.

In the remainder of this section, we study least-cost policy before the constraint binds. We have $\nu(t) = 0$ over these times $t \in [t_0, \tau)$. The least-cost abatement trajectory sets the marginal cost of abatement equal to the shadow cost of CO_2, as given by $\lambda_M(t)$ in equation (4). This is a familiar condition. However, the dynamics of the shadow cost of CO_2 are more interesting than commonly recognized.

First, note that all shadow costs are positive: another unit of temperature or CO_2 requires additional abatement, which raises the cost of the policy program. Using equation (6), the shadow cost of temperature obeys a familiar Hotelling-like condition, adjusted for the effects of climatic inertia:

$$\lambda_T(t) = \lambda_T(t_0)\, e^{(r+\phi)(t-t_0)}. \tag{9}$$

Along an efficient path, the policymaker must be indifferent between accepting another unit of warming in any two instants. The benefit of delaying a unit of warming is composed of the time benefit $r\,\lambda_T(t)$ of delaying the cost by one more instant and also the inertial benefit $\phi\,\lambda_T(t)$ of beginning the following instant with a lower temperature. If there is high inertia (with ϕ small), then temperature would not have changed much between the two instants and the inertial benefit is small. But if there is low inertia (with ϕ large), then temperature would have changed a lot and the inertial benefit is high. Along an efficient path, these benefits must balance the additional cost ($\dot{\lambda}_T(t)$) imposed by delaying the temperature increase. Equating these benefits and costs yields the Hotelling-like condition.

The least-cost abatement policy is determined by the shadow cost of CO_2. From the costate equation (5), the evolution of the least-cost abatement policy is controlled by two terms. The first, positive term is the standard decay-adjusted Hotelling condition familiar from past literature. The second, negative term is novel. Using equations (5) and (9), the appendix shows that the marginal cost of abatement obeys the following relationship along the least-cost trajectory:

(10)

$$\lambda_M(t_0) = e^{-[r+\delta](t-t_0)}\lambda_M(t) + e^{-[r+\delta](t-t_0)}\lambda_T(t) \int_{t_0}^{t} e^{-(\phi-\delta)(t-i)}\phi \, sF'(M(i)) \, di,$$

recalling that $C'(A(t)) = \lambda_M(t)$. The left-hand side is the present cost of abating an additional unit of CO_2 at time t_0. The right-hand side is the present benefit of abating an additional unit of CO_2 at time t_0. The first term is the modified Hotelling term motivated in the introduction and familiar from previous literature. It recognizes that the policymaker should spend fewer dollars early because she discounts future spending and because additional CO_2 emissions have more chance to decay when emitted at an earlier time. If the target were expressed in units of CO_2 rather than temperature, then this would be the only term, and the shadow cost of CO_2 would grow at rate $r + \delta$.[6]

But the target is expressed in units of temperature, not CO_2. The second

[6]This modified Hotelling term is also the only term in the absence of inertia. As $\phi \to \infty$, the integral on the right-hand side of equation (10) vanishes ($\lim_{\phi\to\infty} e^{-\phi(t-i)}\phi = 0$), and because temperature imposes no direct cost without inertia, $\lambda_T(t)$ also vanishes.

component of the present benefit of additional time t_0 abatement describes how it alters time t temperature by changing temperature (via forcing) between times t_0 and t. The total reduction in time t temperature from an additional unit of time t_0 abatement is:[7]

$$\chi(t) \triangleq -\frac{dT(t)}{dA(t_0)} = -\int_{t_0}^{t} \frac{d\dot{T}(i)}{dA(t_0)} di = \int_{t_0}^{t} \left[e^{-\delta(i-t_0)} \phi s F'(M(i)) + \phi \underbrace{\int_{t_0}^{i} \frac{d\dot{T}(j)}{dA(t_0)} dj}_{-\chi(i)} \right] di$$
$$= e^{-\delta(t-t_0)} \int_{t_0}^{t} e^{-(\phi-\delta)(t-i)} \phi s F'(M(i))\, di > 0.$$

The integral describes how additional time t_0 abatement changes time i forcing and how a change in time i forcing changes time t temperature. The present value of the effect of additional time t_0 abatement on time t temperature is $e^{-r(t-t_0)}\lambda_T(t)\chi(t)$, which is the second term on the right-hand side of equation (10).

The appendix develops a phase portrait analysis of the system, establishes further results about the least-cost trajectory, derives least-cost policy under the carbon model of Golosov et al. (2014), and shows that the least-cost trajectory for a "geoengineering" policy has a modified Hotelling form.

III. Calibrated Numerical Example

We now use a calibrated numerical example to estimate the gains from using the least-cost policy program. The appendix gives details and assesses sensitivity to assumptions like constant emissions and geometric decay.

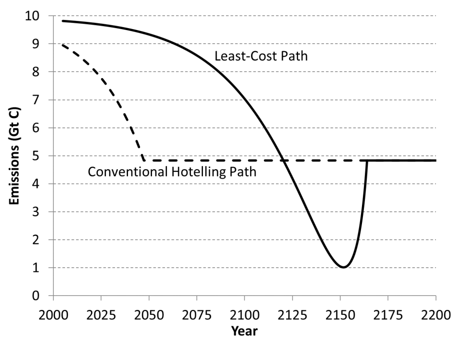
Figure 1 shows how the least-cost path (solid) differs from the standard

[7]The top line uses equation (2) and recognizes that $dM(t)/dA(t_0) = -e^{-\delta(t-t_0)}$. The bottom line follows from converting the top line into a differential equation for $\chi(t)$ and recognizing that $dT(t_0)/dA(t_0) = 0$.

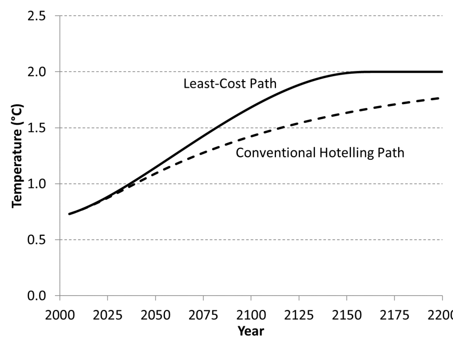
Hotelling solution (dashed), which is the least-cost policy for constraining CO_2 to levels below $\bar{M}$. The climate system's inertia enables the least-cost policy to postpone abatement to later dates without overshooting $\bar{T}$. The Hotelling policy abates emissions too aggressively because it fails to take advantage of the climate system's inertia (top left). Its resulting temperature trajectory is therefore lower than required by the temperature limit (top right), and the system's inertia in fact prevents temperature from ever reaching $\bar{T}$ in finite time under the Hotelling policy. Whereas the least-cost policy overshoots $\bar{M}$ by nearly 100 ppm (bottom left), the Hotelling trajectory never takes advantage of the breathing space afforded by the slowness with which the climate system reacts to overshooting $\bar{M}$. As a consequence, the carbon price starts out much higher under the Hotelling policy and also rises faster until abatement nears its steady-state level (bottom right). However, after the year 2100, the least-cost policy does end up raising the carbon price to levels beyond any reached under the Hotelling trajectory. As CO_2 overshoots its steady-state level, the least-cost policy begins undertaking aggressive abatement so as to reduce CO_2 before temperature exceeds $\bar{T}$. Consistent with the theoretical analysis in the appendix, the efficient carbon price peaks only after CO_2 has peaked, and the carbon price then declines swiftly towards its steady-state value.[8]

The bottom right panel of Figure 1 also plots the Hotelling component (dotted line) of the least-cost carbon price path, as given by the first term

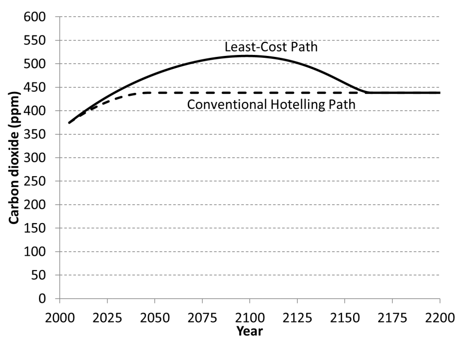
[8]The qualitative properties of our theoretical setting (i.e., CO_2 overshooting its steady-state level and nonmonotonic trajectories for emissions and the carbon price) also appear in temperature-constrained simulations of the benchmark DICE integrated assessment model (Nordhaus, 2008, Chapter 5). Thus, our primary results are robust to including features such as nonstationary business-as-usual emissions, improving abatement technology, savings decisions, pre-threshold damages from temperature change, and more complex carbon and temperature models.



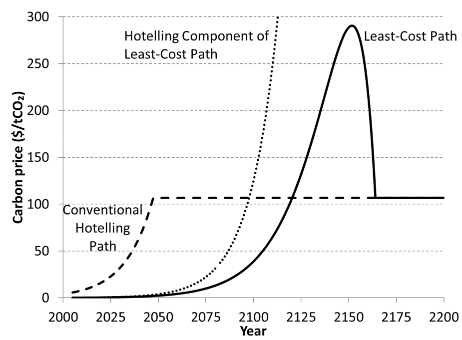
(a) Emissions net of abatement



(b) Temperature



(c) Carbon dioxide



(d) Carbon price

Figure 1. : The least-cost trajectories (solid lines) for emissions, temperature, CO_2, and the carbon price for a temperature limit of $\bar{T} = 2°C$. Also, the conventional Hotelling-like paths (dashed lines), which are also the least-cost paths for the corresponding CO_2 constraint.

in equation (10). Recognizing inertia's braking services makes the least-cost trajectory differ from the Hotelling trajectory in two ways. First, recognizing inertia tends to bend the least-cost trajectory away from its Hotelling component. The gap between the Hotelling component and the least-cost path represents the trajectory adjustment for inertia, which we have seen slows the carbon price's rate of increase. Second, recognizing inertia also reduces the initial carbon price in order to delay abatement. This downward shift in the starting value flattens the Hotelling component of the least-cost trajectory relative to the full Hotelling path (compare the dotted and dashed lines). Near the initial time, the least-cost path differs from the Hotelling path primarily via the downward shift in the initial carbon price. The trajectory adjustment becomes more significant over time, beginning to strongly slow the carbon price's rate of increase near the end of this century, or around the same time that the least-cost CO_2 trajectory peaks.

Table 1 describes how the present cost of the policy program, the year 2005 carbon price, the peak carbon price, and cumulative abatement over the next 200 years vary with the temperature limit $\bar{T}$ and with the recognition of climatic inertia. By ignoring the climate system's inertia, the Hotelling path adds over \$2 trillion in unnecessary costs for a limit of 2°C. Recognizing inertia allows the policymaker to save money both by postponing abatement and by undertaking less cumulative abatement. The climate system's inertia allows for greater natural decay of CO_2 because it delays the temperature consequences of CO_2 emissions (granting more time for decay) and because it allows the CO_2 concentration to overshoot its steady-state level (decay is proportional to the quantity of CO_2). The ability to postpone emission reductions and to undertake fewer emission reductions in total lowers the initial carbon price by over 90%, although the need to bring CO_2 back down to its steady state level increases the peak carbon price by around 200%.[9]

IV. Discussion

We have shown that the least-cost approach to a temperature limit prices carbon emissions at a rate that increases more slowly than exponentially.

[9]For the carbon model of Golosov et al. (2014), the appendix shows that recognizing inertia reduces spending on a 2 degree Celsius temperature limit by nearly \$13 trillion, reduces the initial carbon price by 70%, and increases the peak carbon price by 16%.

Table 1—: The present cost of each policy program, the initial carbon prices, the peak carbon prices, and cumulative abatement over the next 200 years.

	Temperature limit (°C)		
	2	**2.5**	**3**
Cost of efficient path from 2005–2205 ($billion)	98	1.5	0.0001
Cost of Hotelling path from 2005–2205 ($billion)	2,466	181	1.4
CO_2 price along the efficient path in 2005 ($/t$CO_2$)	0.18	0.003	0.000003
CO_2 price along the Hotelling path in 2005 ($/t$CO_2$)	5.8	0.39	0.003
Peak CO_2 price along the efficient path ($/t$CO_2$)	291	164	49
Peak CO_2 price along the Hotelling path ($/t$CO_2$)	107	54	15
Abatement from 2005–2205 along the efficient path (Gt C)	708	266	8
Abatement from 2005–2205 along the Hotelling path (Gt C)	917	540	178

It also temporarily overshoots the steady-state CO_2 level. Computational equilibrium models are the primary tool for estimating the cost of proposed climate policies. These models often assume that the emission price follows an exponential (Hotelling) price path (e.g., Thomson et al., 2011; Bauer et al., 2015) and/or represent a temperature constraint via a constraint on forcing or CO_2 (e.g., Edenhofer et al., 2010). By failing to take advantage of the climate system's inertia, these modeled policies undertake more total abatement than necessary and ramp up policy faster than necessary. Furthermore, these technology-rich integrated assessment models are used to learn about the relative values of prospective low-carbon technologies, but this relative value likely depends on whether the carbon price follows a Hotelling path or instead follows the inverse-U-shaped trajectory described in the present paper. Given that international policy discussions are focused on temperature limits, it should be a high priority to reassess these models' conclusions using frameworks that take advantage of the braking services provided by the climate system's inertia.

REFERENCES

Bauer, Nico, Valentina Bosetti, et al. 2015. "CO2 emission mitigation and fossil fuel markets: Dynamic and international aspects of climate policies." *Technological Forecasting and Social Change*, 90, Part A: 243–256.

Baumol, William J. 1972. "On taxation and the control of externalities." *The American Economic Review*, 62(3): 307–322.

Edenhofer, Ottmar, Brigitte Knopf, et al. 2010. "The economics of low stabilization: Model comparison of mitigation strategies and costs." *Energy Journal*, 31: 11–48.

Gillis, Justin. 2014. "3.6 degrees of uncertainty." *The New York Times.*

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski. 2014. "Optimal taxes on fossil fuel in general equilibrium." *Econometrica*, 82(1): 41–88.

Goulder, Lawrence H., and Koshy Mathai. 2000. "Optimal CO2 abatement in the presence of induced technological change." *Journal of Environmental Economics and Management*, 39(1): 1–38.

Hartl, Richard F., Suresh P. Sethi, and Raymond G. Vickson. 1995. "A survey of the maximum principles for optimal control problems with state constraints." *SIAM Review*, 37(2): 181–218.

Hotelling, Harold. 1931. "The economics of exhaustible resources." *The Journal of Political Economy*, 39(2): 137–175.

Huntingford, Chris, and Jason Lowe. 2007. "“Overshoot” scenarios and climate change." *Science*, 316(5826): 829–829.

Jaeger, Carlo, and Julia Jaeger. 2010. "Three views of two degrees." *Climate Change Economics*, 1(3): 145–166.

Nordhaus, William D. 1980. "Thinking about carbon dioxide: Theoretical and empirical aspects of optimal control strategies." Cowles Foundation Discussion Paper 565.

Nordhaus, William. 1982. "How fast should we graze the global commons?" *The American Economic Review*, 72(2): 242–246.

Nordhaus, William D. 1991. "To slow or not to slow: The economics of the greenhouse effect." *The Economic Journal*, 101(407): 920–937.

Nordhaus, William D. 2008. *A Question of Balance: Weighing the Options on Global Warming Policies.* New Haven:Yale University Press.

Peck, Stephen C., and Y. Steve Wan. 1996. "Analytic solutions of simple optimal greenhouse gas emission models." In *Economics of Atmospheric Pollution.* Vol. 14 of *NATO ASI Series*, ed. Ekko van Ierland and Kazimierz Görka, 113–121. Berlin, Heidelberg:Springer Verlag.

Solomon, Susan, Gian-Kasper Plattner, Reto Knutti, and Pierre Friedlingstein. 2009. "Irreversible climate change due to carbon dioxide emissions." *Proceedings of the National Academy of Sciences*, 106(6): 1704–1709.

Thomson, Allison, Katherine Calvin, et al. 2011. "RCP4.5: a pathway for stabilization of radiative forcing by 2100." *Climatic Change*, 109(1): 77–94.

Tol, Richard S.J. 2013. "Targets for global climate policy: An overview." *Journal of Economic Dynamics and Control*, 37(5): 911–928.

Wetherald, Richard T., Ronald J. Stouffer, and Keith W. Dixon. 2001. "Committed warming and its implications for climate change." *Geophysical Research Letters*, 28(8): 1535–1538.

Wigley, T. M. L., R. G. Richels, and J. A. Edmonds. 2007. "Overshoot pathways to CO2 stabilization in a multi-gas context." In *Human-Induced Climate Change: An Interdisciplinary Assessment*, ed. Michael E. Schlesinger, Haroon Kheshgi, et al., 84–92. New York:Cambridge University Press.

Wigley, Tom M.L. 2003. "Modelling climate change under no-policy and policy emissions pathways." OECD Workshop on the Benefits of Climate Policy: Improving Information for Policy Makers Working Paper ENV/EPOC/GSP(2003)7/FINAL.

Appendix to "Steering the Climate System"

The first section provides additional background on solving control problems with pure state constraints. Section B contains proofs, derivations, and the analysis of the Hotelling-like policy in the main text. Section C describes the numerical example's calibration and solution. Section D adapts the carbon dioxide (CO_2) decay model of Golosov et al. (2014) to our setting and demonstrates that the main text's primary results still hold. Section E allows business-as-usual emissions to evolve over time. Section F numerically explores different degrees of inertia and the role of the discount rate in the base model. Section G provides a phase portrait analysis of the efficient policy. Section H derives the least-cost trajectory for a geoengineering control.

A Optimal control with pure state constraints

We solve our state-constrained control problem via a set of necessary conditions that will look unfamiliar to many economists. The standard approach to solving constrained control problems in economic applications is to embed the Hamiltonian inside of a Lagrangian and apply complementary slackness conditions. This approach requires that a "rank constraint qualification" hold: at any time t, the Jacobian of the binding constraints with respect to the controls must have full rank when evaluated at the optimal control vector $u(t)$ and optimal state vector $x(t)$.[1] Intuitively, the first-order conditions for maximizing a Lagrangian require that the regulator be able to choose its controls so as to have a first-order effect on each binding constraint.

We study a case in which the control $u(t)$ does not enter the constraint (i.e., we have a "pure" state constraint), so that the rank constraint qualification fails to hold. Our time t abatement control can affect a binding time t temperature constraint only by changing temperature at later times. Consider an interval over which a pure state constraint $h(t, x(t), u(t)) \geq 0$ binds. Assume one-dimensional controls and states, and note that being a pure state constraint means $\partial h(t, x(t), u(t))/\partial u(t) = 0$. To maintain the binding constraint $h(t, x(t), u(t)) = 0$, it must be true that $\mathrm{d}h(t, x(t), u(t))/\mathrm{d}t \triangleq h^1(t, x(t), u(t)) = 0$. Maintaining the pure state constraint requires steering the system so that its total derivative with respect to time is 0. If $\partial h^1(t, x(t), u(t))/\partial u(t) = 0$, then maintaining the constraint $h^1(t, x(t), u(t)) = 0$ requires that $\mathrm{d}h^1(t, x(t), u(t))/\mathrm{d}t \triangleq h^2(t, x(t), u(t)) = 0$. We continue this process until finding the first constraint $h^\rho(t, x(t), u(t))$ that includes the control variable $u(t)$. The pure state constraint is then said to be of order ρ. In our setting, the temperature constraint is of order 2 because its first time derivative depends on CO_2 but not abatement, and its second time derivative depends on abatement via the time derivative of CO_2. The policymaker's choice of time t abatement can immediately affect only the time t acceleration

[1] See, for instance, Caputo (2005, Chapter 6).

or deceleration of temperature, not the time t level or velocity of temperature.

Let the pure state constraint be of order ρ. The "indirect adjoining" approach used in our analysis builds the Lagrangian as if $h^\rho(t, x(t), u(t))$ were the relevant constraint. The rank constraint qualification would hold for a system constrained by $h^\rho(t, x(t), u(t)) \geq 0$, but we need to recognize that the true system must actually obey the constraint $h(t, x(t), u(t)) \geq 0$ and, over an interval over which the constraint binds, also $h^k(t, x(t), u(t)) = 0$ for $k < \rho$. Complementary slackness applies to the original constraint $h(t, x(t), u(t))$, not to $h^\rho(t, x(t), u(t))$. Critically, the costate variable on $x(t)$ can jump at the time that the constraint begins to bind.[2] The jump in the costate variable depends on both the partial derivatives of $h^k(t, x(t), u(t))$ with respect to $x(t)$ (for $k < \rho$) and on (the level and time derivatives of) the constraint multiplier.[3]

The survey by Hartl et al. (1995) is the best reference we have found for necessary conditions for control problems with pure state constraints. We adapt the necessary conditions from their Section 6, which presents the indirect adjoining approach to higher-order constraints.

B Formal analysis

This section contains proofs, an additional proposition, the derivation of equation (10), and the analysis of a CO_2 constraint in the main text's setting.

B.1 Proof of Proposition 1

We begin with a lemma that draws on the main text's analysis of the shadow cost of abatement along a least-cost path.

[2]Technically, the costate variable can jump at both the first time that the constraint binds (the "entry time") and the last time that the constraint binds (the "exit time"). However, the values of the costate variable and the constraint multiplier are not unique in that case, so we can normalize the costate variable to jump at only one of the two times. We here choose to allow a jump at the entry time and to impose continuity on the costate variable at the exit time.

[3]Imagine that $\partial h(t, x(t), u(t))/\partial x(t) \geq 0$. Then increasing the state variable helps satisfy the state constraint. Prior to the constraint binding, the costate variable for $x(t)$ includes the value induced by the effect of marginally increasing $x(t)$ on future times' constraints. Intuitively, the costate variable jumps at the time that the constraint begins to bind because the costate variable now includes only the marginal value of the state in meeting $h^\rho(t, x(t), u(t)) \geq 0$; the constraints $h^k(t, x(t), u(t)) = 0$ for $k < \rho$ now do not directly affect the level of the control. In our setting, increasing either CO_2 or (abstracting from a complication due to inertia) temperature makes it more difficult to satisfy the constraint in the future. The shadow costs of CO_2 and temperature initially include these dynamic costs induced by the constraint. Once the constraint begins to bind, the shadow costs of CO_2 and temperature jump down because the temperature constraint now enters the decision problem only as a constraint on the acceleration of temperature.

Lemma B-1. *Let τ indicate the first time $t > t_0$ at which $T(t) = \bar{T}$. Along any least-cost trajectory, τ is finite.*

Proof. The assumption that $E > \delta\left(F^{-1}(\bar{T}/s) - M_{pre}\right)$ implies that temperature along a least-cost path must either reach $\bar{T}$ in finite time or approach it asymptotically from below. Assume that there is no finite time τ at which the system attains $\bar{T}$. Either the system reaches $\bar{M}$ at some finite time and then remains there, or the system approaches $\bar{M}$ asymptotically. Therefore, either abatement reaches $E - \delta\left(F^{-1}(\bar{T}/s) - M_{pre}\right)$ at some finite time and then remains there, or abatement approaches $E - \delta\left(F^{-1}(\bar{T}/s) - M_{pre}\right)$ asymptotically. In either case,

$$\lim_{t\to\infty} \lambda_M(t) = C'\left(E - \delta\left(F^{-1}(\bar{T}/s) - M_{pre}\right)\right), \quad \lim_{t\to\infty} \dot{\lambda}_M(t) = 0.$$

Using equation (5), we have

$$\lim_{t\to\infty} \dot{\lambda}_M(t) = (r+\delta)\, C'\left(E - \delta\left(F^{-1}(\bar{T}/s) - M_{pre}\right)\right) - \phi\, s\, \lambda_T(t_0)\, e^{(r+\phi)(t-t_0)}\, F'(M(t)) = -\infty.$$

But $\dot{\lambda}_M(t)$ cannot approach both zero and negative infinity. We have a contradiction. The time τ must be finite. □

Now consider whether $M(t)$ is greater or less than $\bar{M}$ for $t \in (\tau - \epsilon, \tau)$, for some $\epsilon > 0$.[4] If $M(t) \leq \bar{M}$ for all $t \in (\tau - \epsilon, \tau)$, then equation (2) and $T(t) < \bar{T}$ imply that $T(\tau) = \bar{T}$ only as τ goes to infinity.[5] But Lemma B-1 showed that τ is finite. Therefore $M(t) > \bar{M}$ for some $t \in (\tau - \epsilon, \tau)$. Since this result holds for ϵ arbitrarily small, we then have that there exists some $\Delta > 0$ such that $M(t) > \bar{M}$ for all $t \in (\tau - \Delta, \tau)$.

Once temperature attains $\bar{T}$, CO_2 must remain no larger than $\bar{M}$ in order to prevent temperature from rising past the constraint. And the assumption that $E > \delta\left(F^{-1}(\bar{T}/s) - M_{pre}\right)$ implies that, along a least-cost trajectory, CO_2 must remain no less than $\bar{M}$ once temperature has attained $\bar{T}$. Therefore CO_2 must remain fixed at $\bar{M}$ once temperature attains $\bar{T}$. And because CO_2 must be strictly above $\bar{M}$ at some instant before temperature attains $\bar{T}$, there exists some time q such that $\dot{M}(t) \leq 0$ for all times $t \geq q$ and such that $\dot{M}(t) < 0$ for some time $t \geq q$. This establishes the first part of the proposition.

The second part of the proposition follows immediately from observing that a policymaker constrained to keep CO_2 no greater than the steady-state level $\bar{M}$ corresponding to $\bar{T}$ never lets temperature reach $\bar{T}$. Any path that satisfies the constrained CO_2 problem therefore also satisfies the corresponding constrained temperature problem. However, we have seen that the least-cost CO_2 trajectory must exceed $\bar{M}$ in the constrained temperature problem.

[4] We thank Larry Karp for catching an error in an earlier version of the following proof.

[5] Within the class of trajectories for which $M(t) \leq \bar{M}$ for $t \in (\tau - \epsilon, \tau)$, the trajectory that attains $\bar{T}$ at the earliest time fixes $M(t) = \bar{M}$ for all $t \in (\tau - \epsilon, \tau)$. In that case, equation (2) is an autonomous linear equation, which approaches its steady state $\bar{T}$ only asymptotically. Therefore, if $M(t) \leq \bar{M}$ for $t \in (\tau - \epsilon, \tau)$, then τ must be infinite.

The least-cost path that satisfies the temperature constraint therefore does not satisfy the corresponding CO_2 constraint. Constraining CO_2 introduces an additional binding constraint that strictly increases the cost of the least-cost policy pathway.

B.2 An additional proposition

Proposition B-2. *Let τ be the first time at which $T(t) = \bar{T}$, and let x be the last time prior to τ at which $M(t)$ is nondecreasing. If $\dot{M}(t_0) > 0$, then $x > t_0$, $\dot{\lambda}_M(x) > 0$, and there exists a unique time $y \in (x, \tau)$ at which $\lambda_M(t)$ reaches a maximum.*

Proof. First consider the CO_2 trajectory for times $t \in [t_0, \tau]$. We know by Proposition 1 that it is nonincreasing after some time prior to τ. Combined with the assumption that $\dot{M}(t_0) > 0$, we have that there exists a last time $x \in (t_0, \tau)$ at which $M(t)$ is nondecreasing. At this interior maximum, it must be the case that $\dot{M}(x) = 0$ and $\ddot{M}(x) < 0$. Differentiating equation (1), we have

$$\ddot{M}(t) = -\dot{A}(t) - \delta\, \dot{M}(t).$$

At a point where $\dot{M}(t) = 0$, $\ddot{M}(t) < 0$ if and only if $\dot{A}(t) > 0$. We know by equation (4) that marginal abatement cost equals the shadow cost of CO_2. This establishes that $\dot{\lambda}_M(x) > 0$.

At an interior maximum of $\lambda_M(t)$ in $[t_0, \tau]$, it must be the case that $\dot{\lambda}_M(t) = 0$ and $\ddot{\lambda}_M(t) \leq 0$. Differentiating equation (5), we have:

$$\ddot{\lambda}_M(t) = (r + \delta)\, \dot{\lambda}_M(t) + \Big[-F''(M(t))\, \dot{M}(t) - (r + \phi) F'(M(t))\Big]\, \phi\, s\, \lambda_T(t).$$

At a point where $\dot{\lambda}_M(t) = 0$, $\ddot{\lambda}_M(t) \leq 0$ if and only if $-\dot{M}(t)\, F''(M(t))/F'(M(t)) \leq r + \phi$. Recognizing that $F''(M(t)) < 0$, that $F'(M(t)) > 0$, and that $\dot{M}(t) < 0$ at all times $t \in (x, \tau)$, we have that $\ddot{\lambda}_M(t) < 0$ at any $t \in (x, \tau)$ for which $\dot{\lambda}_M(t) = 0$.

We have already seen that $\dot{\lambda}_M(x) > 0$. Now consider the first time τ when $T(t) = \bar{T}$. The proof of Proposition 1 shows that CO_2 must be strictly greater than $\bar{M}$ in the instants before τ: $M(\tau - \epsilon) = \bar{M} + \gamma$ for ϵ sufficiently small and $\epsilon, \gamma > 0$. In order to achieve the temperature limit at τ, abatement must be such that $[M(\tau) - M(\tau - \epsilon)]/\epsilon = -\gamma/\epsilon$. Letting ϵ and γ jointly go to 0, this relation implies that:

$$\dot{M}(\tau - \epsilon) = E - A(\tau - \epsilon) - \delta(\bar{M} - M_{pre} + \gamma) = -\frac{\gamma}{\epsilon},$$

which holds if and only if:

$$A(\tau - \epsilon) = E - \delta\left(\bar{M} - M_{pre}\right) + \left[\frac{\gamma}{\epsilon} - \delta\, \gamma\right].$$

As ϵ, γ jointly go to 0, the term in the brackets is strictly positive. To maintain temperature at $\bar{T}$ at time τ and beyond, abatement must satisfy $A(\tau) = E - \delta\left(\bar{M} - M_{pre}\right)$. Therefore abatement is greater in the instants before time τ. The main text shows that abatement is continuous at τ. Therefore, $\dot{\lambda}_M(\tau-\epsilon) < 0$. By the Intermediate Value Theorem, there exists some time $y \in (x, \tau-\epsilon)$ such that $\dot{\lambda}_M(y) = 0$. We have already established that $\ddot{\lambda}_M(y) < 0$ for all such y, so there is a unique maximum of $\lambda_M(t)$ between times x and τ.

$\square$

B.3 Derivation of equation (10)

Substitute $\lambda_T(t)$ into equation (5):

$$(r+\delta)\lambda_M(t) - \dot{\lambda}_M(t) = \phi\, sF'(M(t))\,\lambda_T(t_0)\,e^{(r+\phi)(t-t_0)}.$$

Multiply by the integrating factor, integrate with respect to time, and rearrange:

$$\begin{aligned}
&(r+\delta)e^{-(r+\delta)(t-t_0)}\lambda_M(t) - e^{-(r+\delta)(t-t_0)}\dot{\lambda}_M(t) = e^{-(r+\delta)(t-t_0)}\phi\, sF'(M(t))\,\lambda_T(t_0)\,e^{(r+\phi)(t-t_0)}\\
\Leftrightarrow &\int_{t_0}^{t}\left[-(r+\delta)e^{-(r+\delta)(i-t_0)}\lambda_M(i) + e^{-(r+\delta)(i-t_0)}\dot{\lambda}_M(i)\right]\,di\\
&\qquad = \int_{t_0}^{t} -e^{-(r+\delta)(i-t_0)}\phi\, sF'(M(i))\,\lambda_T(t_0)\,e^{(r+\phi)(i-t_0)}\,di\\
\Leftrightarrow &e^{-(r+\delta)(t-t_0)}\lambda_M(t) - \lambda_M(t_0) = -\phi\, s\,\lambda_T(t_0)\int_{t_0}^{t} e^{(\phi-\delta)(i-t_0)}F'(M(i))\,di.
\end{aligned}$$

Substitute in $\lambda_T(t_0) = e^{-(r+\phi)(t-t_0)}\lambda_T(t)$ and rearrange:

$$\lambda_M(t_0) = e^{-[r+\delta](t-t_0)}\lambda_M(t) + e^{-[r+\delta](t-t_0)}\lambda_T(t)\int_{t_0}^{t} e^{-(\phi-\delta)(t-i)}\phi\, sF'(M(i))\,di.$$

B.4 Hotelling policy

Now consider the Hotelling-like policy in the main text's setting. Recall that this policy ignores the inertia in the climate system. It minimizes the cost of meeting the constraint $M(t) \leq \bar{M}$ (while ignoring temperature), where $\bar{M}$ is the unique steady-state CO_2 concentration implied by $\bar{T}$. The Hotelling trajectory solves:

$$\begin{aligned}
&\min_{A(\cdot)} \int_{t_0}^{\infty} e^{-r(t-t_0)}\,C(A(t))\,dt\\
\text{subject to } \dot{M}(t) =& E - A(t) - \delta(M(t)-M_{pre}),\\
M(t) \leq& \bar{M},\\
M(t_0)& \text{ given.}
\end{aligned}$$

We follow the main text in ignoring the nonnegativity constraint on abatement. Define:

$$\begin{aligned} g^0(M(t), A(t)) =& \bar{M} - M(t) \geq 0, \\ g^1(M(t), A(t)) =& -\dot{M}(t). \end{aligned}$$

The state constraint is now of order one. As in other cases we have studied, there will be a first time τ at which the state constraint binds, and the state constraint will then bind forever after τ under a least-cost policy. The current-value Hamiltonian is

$$H(M(t),A(t), \lambda_M(t)) = C(A(t)) + \lambda_M(t)\ [E - A(t) - \delta(M(t) - M_{pre})]\,.$$

The current-value Lagrangian is

$$H[t] + \nu(t)\left\{-E + A(t) + \delta(M(t) - M_{pre})\right\}.$$

The necessary conditions for a maximum are (Hartl et al., 1995):

$$\begin{aligned} C'(A(t)) =& \lambda_M(t) - \nu(t), && \text{(B-1)} \\ \dot{\lambda}_M(t) =& [r + \delta]\, \lambda_M(t) - \nu(t)\, \delta, \\ \nu(t)[\bar{M} - M(t)] =& 0, \quad \nu(t) \leq 0, \quad \dot{\nu}(t) \geq r\, \nu(t), \\ \lambda_M(\tau^-) =& \lambda_M(\tau^+) - e^{r(\tau - t_0)}\, \eta_M, && \text{(B-2)} \\ H[\tau^-] =& H[\tau^+], \\ \eta_M \leq& 0,\ \eta_M \leq e^{-r(\tau - t_0)}\, \nu(\tau^+), && \text{(B-3)} \end{aligned}$$

along with the transition equation, the initial condition on $M(t)$, and the state constraint.[6]

It is easy to see that we get the standard decay-adjusted Hotelling trajectory prior to time τ. After τ, abatement must be chosen so as to hold $\dot{M}(t) = 0$, as in the analysis of a temperature constraint. We need to consider whether abatement jumps at τ. Use equation (B-1) and substitute in from equation (B-2) to obtain:

$$C'(A(\tau^+)) = C'(A(\tau^-)) + e^{r(\tau - t_0)}\, \eta_M - \nu(\tau^+).$$

The conditions in (B-3) then imply that abatement either jumps down at τ (if $\eta_M < e^{-r(\tau - t_0)}\, \nu(\tau^+)$) or is continuous at τ (if $\eta_M = e^{-r(\tau - t_0)}\, \nu(\tau^+)$). Assume that abatement

[6] Formally, there are two more necessary conditions: that $\mathrm{d}H[t]/\mathrm{d}t = \mathrm{d}L[t]/\mathrm{d}t$, and that an omitted multiplier on the instantaneous payoff function be weakly positive. The first condition is always satisfied since at any time t either the constraints are binding or their Lagrange multipliers are zero. The second condition cannot be satisfied with a multiplier of zero because abatement would then always be either at its upper or lower bound (in order to maximize the Lagrangian), which cannot be optimal. Thus, as is typical in economic analysis, the omitted multiplier must be strictly positive and therefore ignorable. For ease of presentation, we ignore these two conditions here, in the main text, and in the remainder of the appendix.

jumps down at τ. We then have:

$$\begin{aligned}\dot{M}(\tau^-) =& E - A(\tau^-) - \delta(M(\tau^-) - M_{pre})\\ <& E - A(\tau^+) - \delta(M(\tau^-) - M_{pre})\\ =& E - A(\tau^+) - \delta(M(\tau^+) - M_{pre})\\ =& \dot{M}(\tau^+)\\ =& 0.\end{aligned}$$

Therefore, if abatement jumps down at τ, then $\dot{M}(\tau^-) < 0$, which would imply that CO_2 is declining towards $\bar{M}$ and thus that $M(t) > \bar{M}$ for some time $t < \tau$. But this would violate the state constraint. We have a contradiction. As a result, abatement must be continuous at τ and we must have $\eta_M = e^{-r(\tau - t_0)}\,\nu(\tau^+)$.

C Numerical calibration and solution

We calibrate the example to DICE-2007 (Nordhaus, 2008), as implemented with an annual timestep in Lemoine and Traeger (2014). All baseline runs use the 5.5% annual consumption discount rate ($r = 0.055$) generally consistent with this model.[7]

The full DICE model includes three carbon reservoirs. Lemoine and Traeger (2014) approximate DICE's full carbon dynamics by making the decay rate of CO_2 a function of the atmospheric CO_2 stock and time. Along the optimal path in DICE, the time-varying decay rate for CO_2 in excess of its pre-industrial level starts at 0.0141, declines to 0.0119 in 100 years, and declines to 0.0068 after 200 years. Using the average value over the first 100 years, we have $\delta = 0.0138$. We calibrate business-as-usual CO_2 emissions E to DICE's initial value. This yields $E = 9.97$ Gt C per year.

We follow much scientific literature in modeling forcing as $F(M(t)) = \alpha \ln(M(t)/M_{pre})$. We take $M_{pre} = 596.4$ Gt C, and we follow Ramaswamy et al. (2001, Table 6.2) in using $\alpha = 5.35$ W m^{-2}, which is approximately equivalent to the parameters used in DICE. The full DICE model includes two temperature reservoirs. Lemoine and Traeger (2014) simplify this setting by representing the deep ocean temperature as a function $\gamma_T(T, t)$ of surface temperature and time. In their discrete-time setting, the temperature transition becomes

$$T_{t+1} - T_t = C_T \left[F_{t+1} - \frac{\alpha \ln(2)}{cs} T_t - [1 - \gamma_T(T_t, t)]\, C_O\, T_t \right],$$

where we have used cs for climate sensitivity so as to avoid confusion with the present paper's notation. The present paper's parameter s gives equilibrium warming per unit of forcing,

[7] Technically, this setting with stationary output should use a discount rate no greater than 1.5% to be consistent with DICE-2007: consumption growth in the Ramsey equation is negative once we subtract the cost of abatement.

whereas DICE's $cs = 3$ gives equilibrium warming from doubled CO_2. Relating the two parameters, we have:

$$s = \frac{cs}{\alpha \ln(2)} = 0.809 \text{ °C } \left[\text{W m}^{-2}\right]^{-1}.$$

Using explicit Euler difference methods, we find:

$$\phi = \frac{C_T \left[F_{t+1} - \frac{\alpha \ln(2)}{cs} T_t - \left[1 - \gamma_T(T_t, t)\right] C_O \, T_t\right]}{s \, F_t - T_t}.$$

Along DICE's optimal trajectory, the inferred value of ϕ starts at 0.0129, falls to 0.0056 after 100 years, and falls to -0.0030 after 200 years (reflecting that the ocean begins transferring heat to the atmosphere as the CO_2 concentration declines). Using the average value over the first 100 years, we have $\phi = 0.0091$.

In DICE, the cost (as a fraction of time t output) of abating a fraction μ_t of business-as-usual emissions is $\Psi_t \mu_t^{a_2}$, where $a_2 = 2.8$ and

$$\Psi_t = \frac{a_0 \, \sigma_t}{a_2} \left(1 - \frac{1 - e^{(t - t_0) \, g_\Psi}}{a_1}\right), \text{ with } \sigma_t = \sigma_0 \exp\left[\frac{g_{\sigma,0}}{\delta_\sigma}\left(1 - e^{-(t - t_0) \, \delta_\sigma}\right)\right].$$

The parameters are $a_0 = 1.17$, $a_1 = 2$, $g_\Psi = -0.005$, $\sigma_0 = 0.13$, $g_{\sigma,0} = -0.0073$, and $\delta_\sigma = 0.003$. Initial output Y (without adjusting for climate damages) in DICE is approximately 85 trillion dollars. We represent the cost of abatement $A(t)$ as

$$C\left(A(t)\right) = \Psi_{t_0} \left[\frac{A(t)}{E}\right]^{a_2} Y. \tag{C-4}$$

Finally, from DICE-2007, we have the initial CO_2 stock as $M_0 = 808.9$ Gt C, the initial global mean surface temperature as $T_0 = 0.7307$ °C (relative to 1900), and the initial time as $t_0 = 2005$.

To solve the four-dimensional system of differential equations defined in the main text, we begin with a triplet $(T(\tau), M(\tau), A(\tau))$ such that $T(\tau) = \bar{T}$, $M(\tau) = \bar{M}$, and $A(\tau)$ holds $\dot{M}(\tau) = 0$. From the Maximum Principle, we have $\lambda_M(\tau^-) = C'(A(\tau))$. We then seek the value of $\lambda_T(\tau^-)$ consistent with these conditions and with the initial conditions. For a given value of $\lambda_T(\tau^-)$, we use Matlab's ode23 solver with relative and absolute tolerances of 10^{-10} to solve the system of ordinary differential equations from τ but with time flowing in reverse.[8] In the resulting simulation, let x be the time t at which $M(t) = M_0$. At a solution to the system, it must also be the case that $T(x) = T_0$. An optimization routine searches for

[8]In general, we cannot solve the model forward by searching for the initial shadow costs $\lambda_M(t_0)$ and $\lambda_T(t_0)$ that lead the system to obey the terminal conditions because, as is typical of saddle-path stable systems, values slightly off the desired trajectories lead the system to a wildly different outcome. Our solution method is closely related to the "reverse shooting" technique described in Judd (1998).

the value of $\lambda_T(\tau^-)$ such that $T(x) = T_0$. At a solution, the values $\lambda_M(x)$ and $\lambda_T(x)$ are the efficient $\lambda_M(t_0)$ and $\lambda_T(t_0)$.[9] Using these initial values, we then simulate the model forward in actual time, setting $\lambda_M(t)$ to hold $M(t)$ constant at $M(\tau)$ for all times $t > \tau$. We use the trapezoidal method to approximate the integral of abatement cost over the mesh points.

D Extension to the decay model of Golosov et al. (2014)

Our main analysis assumes that the stock of CO_2 decays exponentially. In reality, the evolution of atmospheric CO_2 is more complex. We here extend our setting to the more realistic decay model of Golosov et al. (2014).

In Golosov et al. (2014), a fraction ψ_L of emissions remains forever, a fraction $(1-\psi_0)(1-\psi_L)$ decays immediately, and a fraction $\psi_0(1-\psi_L)$ decays geometrically at rate ψ. Their carbon decay model reduces to the main text's model when $\psi_L = 0$, $\psi_0 = 1$, and $\psi = \delta$. Let $M_1(t)$ be the stock of CO_2 that remains in the atmosphere forever and $M_2(t)$ be the stock of CO_2 that decays geometrically. We have the following equations of motion:

$$\begin{aligned}\dot{M}_1(t) =& \psi_L[E - A(t)],\\ \dot{M}_2(t) =& \psi_0(1-\psi_L)[E - A(t)] - \psi M_2(t).\end{aligned}$$

The total stock of CO_2 is the sum of the CO_2 in these two atmospheric reservoirs: $M(t) = M_1(t) + M_2(t)$. When we numerically implement this model, we follow their calibration in using $M_1(t_0) = 684$ Gt C, $M_2(t_0) = 118$ Gt C, $\psi_L = 0.2$, $\psi_0 = 0.393$, and $\psi = 0.0228/10$, where the latter adjusts for measuring time in years rather than in decades.[10]

The next subsection analyzes the least-cost policy trajectory with this new decay model, the subsequent subsection analyzes the least-cost Hotelling trajectory, and the third subsection reports numerical results.

D.1 Least-cost policy

We now consider least-cost policy. As in the main analysis, the CO_2 stock must equal $\bar{M}$ when the policymaker decides to finally let $T(t)$ reach $\bar{T}$, because otherwise the constraint $T(t) \leq \bar{T}$ either would be violated or would fail to bind in the following instant. As before, let τ be the first time at which $T(t) = \bar{T}$. The efficient policy trajectory must keep $M(t) = \bar{M}$ for all $t \geq \tau$.

[9] When solving for the Hotelling trajectory, we begin with $M(\tau)$ equal to $\bar{M}$ and $\lambda_M(\tau^-)$ equal to $C'(A(\tau))$. No search is necessary, as temperature can be effectively removed from the policymaker's problem.

[10] Golosov et al. (2014) abstract from inertia: they assume that temperature responds instantly to CO_2. They describe how to adjust their carbon decay model to mimic the combined effects of thermal inertia and carbon decay in the DICE model. We do not use this adjustment because we model inertia explicitly and we want to analyze robustness to their own carbon decay model.

Let $M_{min}(t)$ be the minimum CO_2 stock attainable at any time after t. If emissions (net of abatement) are strictly positive, then the infinite persistence of a fraction ψ_L of CO_2 means that $M_{min}(t)$ is increasing over time ($E(t) - A(t) > 0 \Leftrightarrow \dot{M}_{min}(t) > 0$). For any finite time t, the efficient policy path requires that $M_{min}(t) < \bar{M}$ so that the temperature constraint will not be violated at some later time. In order to prevent $M_{min}(t)$ from eventually becoming larger than $\bar{M}$, the efficient policy must have $A(t) \to E$ (at which point $\dot{M}_{min}(t) = 0$). We have thus seen that the policymaker must eventually eliminate all emissions. This result contrasts with the main text's setting with geometric decay, in which strictly positive emissions are consistent with holding the CO_2 stock fixed at $\bar{M}$ for all times $t \geq \tau$.[11]

We can also see that the policymaker eliminates all emissions only asymptotically. Imagine that the optimal path is such that $A(t) = E$ for finite t. At that time, the total stock of CO_2 would be declining because of the geometric decay represented by ψ. But this declining stock violates the condition that an efficient trajectory holds $M(t)$ fixed at $\bar{M}$ for all times $t \geq \tau$. For t sufficiently large, the efficient trajectory must have $A(t) \to E$ only asymptotically.

The least-cost abatement trajectory must solve:

$$\begin{aligned}
&\min_{A(\cdot)} \int_{t_0}^{\infty} e^{-r(t-t_0)}\, C(A(t))\, dt \\
\text{subject to } \dot{M}_1(t) =&\psi_L[E - A(t)], \\
\dot{M}_2(t) =&\psi_0(1-\psi_L)[E - A(t)] - \psi M_2(t), \\
\dot{T}(t) =&\phi\left[s\, F(M_1(t) + M_2(t)) - T(t)\right], \\
A(t) \leq& E, \\
T(t) \leq& \bar{T}, \\
&M_1(t_0), M_2(t_0), T(t_0) \text{ given.}
\end{aligned}$$

In contrast to the main text, we explicitly represent the nonnegativity constraint on net emissions $E - A(t)$. We will see in the numerical analysis that the new decay model makes this constraint relevant. Following the main text, define:

$$\begin{aligned}
h^0(T(t), M_1(t), M_2(t), A(t)) =&\bar{T} - T(t) \geq 0, \\
h^1(T(t), M_1(t), M_2(t), A(t)) =& - \dot{T} = -\phi\left[s\, F(M_1(t) + M_2(t)) - T(t)\right], \\
h^2(T(t), M_1(t), M_2(t), A(t)) =& - \ddot{T} = -\phi\left[s\, F'\left(M_1(t) + M_2(t)\right)\left(\dot{M}_1(t) + \dot{M}_2(t)\right) - \dot{T}(t)\right] \\
=& -\phi s\, F'(M_1(t) + M_2(t))\left\{[\psi_0(1-\psi_L) + \psi_L][E - A(t)] - \psi M_2(t)\right\} \\
&+ \phi^2\left[s\, F(M_1(t) + M_2(t)) - T(t)\right].
\end{aligned}$$

[11]Further, because a fraction of emissions persists forever, the temperature limit here fixes cumulative emissions. In contrast, in the main text, we saw that recognizing inertia enabled the policymaker to increase cumulative emissions.

As in the main text, the state constraint is of order two. The current-value Hamiltonian is

$$\begin{aligned}H(M_1(t),&M_2(t),T(t),A(t),\lambda_{M1}(t),\lambda_{M2}(t),\lambda_T(t))\\&=C(A(t))+\lambda_{M1}(t)\,\psi_L[E-A(t)]+\lambda_{M2}(t)\;\{\psi_0[1-\psi_L][E-A(t)]-\psi M_2(t)\}\\&\quad+\lambda_T(t)\,\phi\,[s\,F(M_1(t)+M_2(t))-T(t)].\end{aligned}$$

The current-value Lagrangian is

$$\begin{aligned}&H[t]+\mu(t)\,[A(t)-E]\\&+\nu(t)\Big\{-\phi s\,F'(M_1(t)+M_2(t))\,\{[\psi_0(1-\psi_L)+\psi_L][E-A(t)]-\psi M_2(t)\}\\&\qquad+\phi^2\;[s\,F(M_1(t)+M_2(t))-T(t)]\Big\}.\end{aligned}$$

The necessary conditions for a maximum are (Hartl et al., 1995):

$$\begin{aligned}C'(A(t))=&\lambda_{M1}(t)\psi_L+\lambda_{M2}(t)\psi_0[1-\psi_L]-\mu(t)\\&-\nu(t)\phi\,s\,F'(M_1(t)+M_2(t))[\psi_0(1-\psi_L)+\psi_L], && \text{(D-5)}\\ \dot\lambda_{M1}(t)=&r\lambda_{M1}(t)-\phi\,s\,F'(M_1(t)+M_2(t))\,\lambda_T(t)\\&+\nu(t)\,\phi\,s\,F''(M_1(t)+M_2(t))\,\{[\psi_0\,(1-\psi_L)+\psi_L][E-A(t)]-\psi M_2(t)\}\\&-\nu(t)\,\phi^2\,s\,F'(M_1(t)+M_2(t)),\\ \dot\lambda_{M2}(t)=&[r+\psi]\lambda_{M2}(t)-\phi\,s\,F'(M_1(t)+M_2(t))\,\lambda_T(t)\\&+\nu(t)\,\phi\,s\,F''(M_1(t)+M_2(t))\,\{[\psi_0\,(1-\psi_L)+\psi_L][E-A(t)]-\psi M_2(t)\}\\&-\nu(t)\,[\phi+\psi]\phi\,s\,F'(M_1(t)+M_2(t)),\\ \dot\lambda_T(t)=&[r+\phi]\lambda_T(t)+\nu(t)\phi^2,\\ \mu(t)\ge&0,\;\;A(t)-E\le 0,\;\;\mu(t)\,[A(t)-E]=0,\\ \nu(t)[\bar T-&T(t)]=0,\quad \nu(t)\le 0,\quad \dot\nu(t)\ge r\,\nu(t),\quad \ddot\nu(t)\le 2\,r\dot\nu(t)-r^2\,\nu(t),\\ \lambda_{M1}(\tau^-)=&\lambda_{M1}(\tau^+)-e^{r(\tau-t_0)}\,\eta^2_{M1}\phi\,s\,F'(M_1(t)+M_2(t)), && \text{(D-6)}\\ \lambda_{M2}(\tau^-)=&\lambda_{M2}(\tau^+)-e^{r(\tau-t_0)}\,\eta^2_{M2}\phi\,s\,F'(M_1(t)+M_2(t)), && \text{(D-7)}\\ \lambda_T(\tau^-)=&\lambda_T(\tau^+)-e^{r(\tau-t_0)}\,\eta^1_T+e^{r(\tau-t_0)}\,\eta^2_T\phi,\\ H[\tau^-]=&H[\tau^+],\end{aligned}$$

$$\eta^1_x,\eta^2_x\le 0,\;\;\eta^1_x\le -\,e^{-r(\tau-t_0)}\,\dot\nu(\tau^+)+re^{-r(\tau-t_0)}\nu(\tau^+),\;\;\eta^2_x=e^{-r(\tau-t_0)}\,\nu(\tau^+)\;\;\text{for } x\in\{T,M1,M2\}, \qquad \text{(D-8)}$$

along with the transition equations, the initial conditions on $M_1(t)$, $M_2(t)$, and $T(t)$, and the state constraint.

For times $t \geq \tau$, abatement evolves so as to keep $h^1 = 0$, which requires $M_1(t) + M_2(t) = \bar{M}$. In order to maintain $h^2 = 0$ (i.e., in order to stay at $\bar{M}$), we need:

$$A(t) = E - \frac{\psi}{\psi_L + \psi_0(1-\psi_L)} \left[\bar{M} - M_1(t)\right], \tag{D-9}$$

for $t \geq \tau$. We therefore have:

$$A(\tau) = E - \frac{\psi}{\psi_L + \psi_0(1-\psi_L)} \left[\bar{M} - M_1(\tau)\right] \triangleq \bar{A}(M_1(\tau)).$$

Differentiate equation (D-9) with respect to time:

$$\dot{A}(t) = \frac{\psi\,\psi_L}{\psi_L + \psi_0(1-\psi_L)} [E - A(t)] \geq 0,$$

for $t \geq \tau$. Integrating from τ to $t \geq \tau$ yields:

$$A(t-\tau; M_1(\tau)) = E - e^{-\frac{\psi\,\psi_L}{\psi_L+\psi_0(1-\psi_L)}(t-\tau)}[E - \bar{A}(M_1(\tau))].$$

After temperature reaches $\bar{T}$, abatement rises from $A(\tau) = \bar{A}(M_1(\tau))$ towards E and attains E only asymptotically, as argued above.

From equation (D-5), we have

$$C'(A(\tau^-)) = \lambda_{M1}(\tau^-)\psi_L + \lambda_{M2}(\tau^-)\psi_0[1-\psi_L] - \mu(\tau^-),$$

and also that

$$C'(A(\tau^+)) = \lambda_{M1}(\tau^+)\psi_L + \lambda_{M2}(\tau^+)\psi_0[1-\psi_L] - \mu(\tau^+) - \nu(\tau^+)\phi\, s\, F'(M_1(\tau)+M_2(\tau))[\psi_0(1-\psi_L)+\psi_L].$$

Equations (D-6) and (D-7) and the conditions in (D-8) then imply that $A(\tau^-) = A(\tau^+)$. Thus, as in the main text, abatement is continuous at time τ.

In the remainder of this section, we study least-cost policy before the constraint binds. We have $\nu(t) = 0$ over these times $t \in [t_0, \tau)$. At these times, $C'(A(t)) = \lambda_{M1}(t)\psi_L + \lambda_{M2}(t)\psi_0[1-\psi_L] - \mu(t)$. If $\mu(t) > 0$, then $A(t) = E$ and $\mu(t)$ picks up the gap between the shadow cost of emissions and $C'(E)$ in equation (D-5).

Following the derivation for the main text, the costate equations on $M_1(t)$, $M_2(t)$, and $T(t)$ imply the following relationships:

$$\begin{aligned}
\lambda_{M1}(t_0) =& e^{-r(t-t_0)}\,\lambda_{M1}(t) + e^{-r(t-t_0)}\,\lambda_T(t) \int_{t_0}^{t} e^{-\phi(t-i)}\,\phi\, s\, F'(M(i))\,di, \\
\lambda_{M2}(t_0) =& e^{-[r+\psi](t-t_0)}\,\lambda_{M2}(t) + e^{-[r+\psi](t-t_0)}\,\lambda_T(t) \int_{t_0}^{t} e^{-(\phi-\psi)(t-i)}\,\phi\, s\, F'(M(i))\,di.
\end{aligned}$$

These equations are exactly the same as in the main text, except with $M_1(t)$ lacking a geometric decay component. We therefore see that inertia enters $\lambda_{M1}(t)$ and $\lambda_{M2}(t)$ in the same way as it entered $\lambda_M(t)$ in the main text (modulo the geometric decay terms in the exponents). Further, equation (D-5) shows that marginal abatement cost (which defines the efficient emission tax) is linear in $\lambda_{M1}(t)$ and $\lambda_{M2}(t)$, just as it was linear in $\lambda_M(t)$ in the main text. The way that inertia affects the efficient tax on emissions is therefore qualitatively unchanged by the extension to the more realistic decay model of Golosov et al. (2014).

Write the cost of the remaining policy program at τ as a function of τ and $M_1(\tau)$:

$$W(\tau, M_1(\tau)) = \int_\tau^\infty e^{-r(i-\tau)} C(A(i-\tau; M_1(\tau)))\,\mathrm{d}i.$$

Along a least-cost path, the costate variables must be[12]

$$\lambda_{M1}(\tau^+) = \frac{\partial W(\tau, M_1(\tau))}{\partial M_1(\tau)} = \frac{\psi}{\psi_L + \psi_0(1-\psi_L)} \int_\tau^\infty e^{-\left(r+\frac{\psi\,\psi_L}{\psi_L+\psi_0(1-\psi_L)}\right)(i-\tau)} C'(A(i-\tau; M_1(\tau)))\,\mathrm{d}i$$

and

$$\lambda_{M2}(\tau^+) = \frac{\partial W(\tau, M_1(\tau))}{\partial M_2(\tau)} = 0.$$

From equations (D-6) and (D-7) and the conditions in (D-8), we then have:

$$\begin{aligned}\lambda_{M1}(\tau^-) =& \frac{\psi}{\psi_L + \psi_0(1-\psi_L)} \int_\tau^\infty e^{-\left(r+\frac{\psi\,\psi_L}{\psi_L+\psi_0(1-\psi_L)}\right)(i-\tau)} C'(A(i-\tau; M_1(\tau)))\,\mathrm{d}i \\ & - \nu(\tau^+)\,\phi\,s\,F'(\bar{M}), && \text{(D-10)} \\ \lambda_{M2}(\tau^-) =& - \nu(\tau^+)\,\phi\,s\,F'(\bar{M}). && \text{(D-11)}\end{aligned}$$

From equation (D-5), we then have:

$$\begin{aligned}C'(\bar{A}(M_1(\tau))) =& \psi_L \frac{\psi}{\psi_L + \psi_0(1-\psi_L)} \int_\tau^\infty e^{-\left(r+\frac{\psi\,\psi_L}{\psi_L+\psi_0(1-\psi_L)}\right)(i-\tau)} C'(A(i-\tau; M_1(\tau)))\,\mathrm{d}i \\ & - [\psi_L + \psi_0[1-\psi_L]]\,\nu(\tau^+)\,\phi\,s\,F'(\bar{M}), && \text{(D-12)}\end{aligned}$$

where we recognize that the abatement nonnegativity constraint does not bind at τ. (Suppose the constraint did bind at τ. We know that $M_2(\tau) > 0$, resulting in $M_1 + M_2 < \bar{M}$ in the

[12]If we had instead defined $\bar{A}$ as a function of $M_2(\tau)$ and left $M_1(\tau)$ as the residual, then we would obtain $\lambda_{M1}(\tau^+) = 0$ and $\lambda_{M2}(\tau^+) < 0$, with the difference between them being the exact same as in the given analysis. We will see that it is the difference that matters, as $\nu(\tau)$ will work to shift both multipliers' right-hand limits to match their left-hand limits. The results needed for the numerics will therefore be unchanged, as the inferred value of $\nu(\tau^+)$ will simply reflect whichever choice we make. However, the given presentation with both costate variables positive matches the reasonable assumption that the shadow costs should be positive.

next instant, but the efficient policy must maintain CO_2 at $\bar{M}$. The abatement nonnegativity constraint therefore cannot bind at τ.)

To numerically solve the model, we guess $\lambda_T(\tau^-)$ and $M_1(\tau)$. The guess for $M_1(\tau)$ implies $A(\tau)$, which in turn implies $\nu(\tau^+)$ from equation (D-12) and then $\lambda_{M1}(\tau^-)$ and $\lambda_{M2}(\tau^-)$ from equations (D-10) and (D-11).[13] We also know that $T(\tau) = \bar{T}$ and $M_2(\tau) = \bar{M} - M_1(\tau)$. We solve the system backwards in time, stopping when either $M_1(t)$ meets its initial condition or $A(t) = E$. In the latter case, we then simulate the system backwards from this new point, with $\mu(t)$ starting at zero (its value at the latest time that the constraint binds).[14] The differential equation for $\mu(t)$ comes from fixing $A(t) = E$ and then differentiating equation (D-5) with respect to time. We simulate the system backwards in time with $A(t) = E$ until we find a time at which $\mu(t)$ once again reaches zero, which is where the constraint that $E - A(t) \geq 0$ just started to bind. We then simulate the unconstrained system backwards in time from there, stopping when $M_1(t)$ meets its initial condition. Once we have found a time when $M_1(t)$ meets its initial condition, we check the initial conditions on $M_2(t)$ and $T(t)$. We iterate until we find values of $\lambda_T(\tau^-)$ and $M_1(\tau)$ that generate paths that satisfy these initial conditions.

D.2 Hotelling policy

Now consider the Hotelling-like policy under the decay model of Golosov et al. (2014). Recall that this policy ignores the inertia in the climate system. In the main analysis, it minimizes the cost of meeting the constraint $M(t) \leq \bar{M}$ (while ignoring temperature), where $\bar{M}$ is the unique steady-state CO_2 concentration implied by $\bar{T}$. Here, we study the problem of constraining $M_1(t) + M_2(t) \leq \bar{M}$, while ignoring temperature. The Hotelling trajectory solves:

$$\begin{aligned}
&\min_{A(\cdot)} \int_{t_0}^{\infty} e^{-r(t-t_0)}\, C(A(t))\, dt \\
\text{subject to } \dot{M}_1(t) &= \psi_L[E - A(t)], \\
\dot{M}_2(t) &= \psi_0(1-\psi_L)[E - A(t)] - \psi M_2(t), \\
A(t) &\leq E, \\
M_1(t) + M_2(t) &\leq \bar{M}, \\
&M_1(t_0), M_2(t_0) \text{ given.}
\end{aligned}$$

[13] We approximate the integral in equation (D-12) by starting from near $M_1(t) = \bar{M}$ and $M_2(t) = 0$, simulating backwards until reaching $M_1(\tau)$, and then using a Newton-Cotes formula to approximate the integral.

[14] In general, $\mu(t)$ need be only piecewise continuous, but continuity of $A(t)$ here ensures continuity of $\mu(t)$. See Caputo (2005, Chapter 6).

Define:

$$\begin{aligned} g^0(M_1(t), M_2(t), A(t)) =& \bar{M} - M_1(t) - M_2(t) \geq 0, \\ g^1(M_1(t), M_2(t), A(t)) =& - \dot{M}_1(t) - \dot{M}_2(t). \end{aligned}$$

The state constraint is now of order one. As in other cases we have studied, there will be a first time τ at which the state constraint binds, and it will bind forever after that time under a least-cost policy. The current-value Hamiltonian is

$$\begin{aligned} H(M_1(t),& M_2(t), A(t), \lambda_{M1}(t), \lambda_{M2}(t)) \\ &= C(A(t)) + \lambda_{M1}(t)\, \psi_L [E - A(t)] + \lambda_{M2}(t)\, \{\psi_0 [1 - \psi_L][E - A(t)] - \psi M_2(t)\} . \end{aligned}$$

The current-value Lagrangian is

$$H[t] + \mu(t)\, [A(t) - E] + \nu(t)\, \{-[\psi_0(1 - \psi_L) + \psi_L][E - A(t)] + \psi M_2(t)\} .$$

The necessary conditions for a maximum are (Hartl et al., 1995):

$$\begin{aligned} C'(A(t)) =& \lambda_{M1}(t)\psi_L + \lambda_{M2}(t)\psi_0[1 - \psi_L] - \mu(t) - \nu(t)\, [\psi_0(1 - \psi_L) + \psi_L], && \text{(D-13)} \\ \dot{\lambda}_{M1}(t) =& r\lambda_{M1}(t), \\ \dot{\lambda}_{M2}(t) =& [r + \psi]\lambda_{M2}(t) - \nu(t)\, \psi, \\ \mu(t) \geq& 0,\ \ A(t) - E \leq 0,\ \ \mu(t)\, [A(t) - E] = 0, \\ \nu(t)[\bar{M}& - M_1(t) - M_2(t)] = 0, \quad \nu(t) \leq 0, \quad \dot{\nu}(t) \geq r\, \nu(t), \\ \lambda_{M1}(\tau^-) =& \lambda_{M1}(\tau^+) - e^{r(\tau - t_0)}\, \eta_{M1}, && \text{(D-14)} \\ \lambda_{M2}(\tau^-) =& \lambda_{M2}(\tau^+) - e^{r(\tau - t_0)}\, \eta_{M2}, && \text{(D-15)} \\ H[\tau^-] =& H[\tau^+], \\ \eta_x \leq& 0,\ \ \eta_x \leq e^{-r(\tau - t_0)}\, \nu(\tau^+) \ \text{ for } x \in \{M1, M2\}, && \text{(D-16)} \end{aligned}$$

along with the transition equations, the initial conditions on $M_1(t)$ and $M_2(t)$, and the state constraint.

For times $t \geq \tau$, abatement evolves so as to keep $g^1 = 0$. This requirement generates the same post-τ policy path as in the previous subsection. Now consider whether abatement is continuous at τ. Use equation (D-13) and substitute in from equations (D-14) and (D-15) to obtain:

$$C'(A(\tau^+)) = C'(A(\tau^-)) + e^{r(\tau - t_0)}\, [\psi_L\, \eta_{M1} + \psi_0\, (1 - \psi_L)\, \eta_{M2}] - (\psi_0(1 - \psi_L) + \psi_L)\, \nu(\tau^+).$$

The conditions in (D-16) then imply that abatement either jumps down at τ (if either $\eta_{M1} < e^{-r(\tau - t_0)}\, \nu(\tau^+)$ or $\eta_{M2} < e^{-r(\tau - t_0)}\, \nu(\tau^+)$) or is continuous at τ (if $\eta_{M1} = \eta_{M2} =$

$e^{-r(\tau-t_0)}\,\nu(\tau^+))$. Assume that abatement jumps down at τ. We then have:

$$\begin{aligned}\dot{M}_1(\tau^-)+\dot{M}_2(\tau^-) =&[\psi_L+\psi_0(1-\psi_L)][E-A(\tau^-)]-\psi M_2(\tau^-)\\ <&[\psi_L+\psi_0(1-\psi_L)][E-A(\tau^+)]-\psi M_2(\tau^-)\\ =&[\psi_L+\psi_0(1-\psi_L)][E-A(\tau^+)]-\psi M_2(\tau^+)\\ =&\dot{M}_1(\tau^+)+\dot{M}_2(\tau^+)\\ =&0.\end{aligned}$$

Therefore, if abatement jumps down at τ, then $\dot{M}_1(\tau^-)+\dot{M}_2(\tau^-)<0$, which would imply that total CO_2 is declining towards $\bar{M}$ and thus that $M_1(t)+M_2(t)>\bar{M}$ for some time $t<\tau$. But this would violate the state constraint. We have a contradiction. As a result, abatement must be continuous at τ and $\eta_{M1}=\eta_{M2}=e^{-r(\tau-t_0)}\,\nu(\tau^+)$.

The remaining analysis and the numerical methods are directly analogous to the previous subsection. Note that each shadow cost increases exponentially for $t\in(t_0,\tau)$. We therefore recover a Hotelling-like trajectory, modified for this decay model.

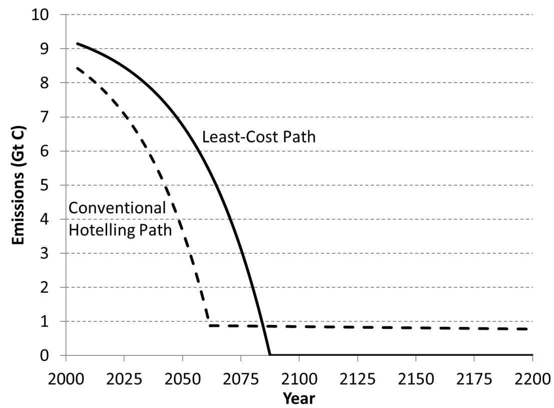
D.3 Numerical example

We now extend the numerical example from the main text to the decay model of Golosov et al. (2014). Figure D1 depicts the least-cost paths for emissions, temperature, each stock of carbon dioxide, and the emission tax implied by a 2 degree Celsius temperature constraint, along with the "Hotelling" paths generated by constraining $M(t)\leq\bar{M}$. This figure is the analogue of Figure 1 in the main text. As in the main text, we see that the Hotelling policy reduces emissions more aggressively than does the least-cost policy.[15] Temperature therefore increases more slowly under the Hotelling tax trajectory and only asymptotically approaches the constraint $\bar{T}$ (top right). As expected, the new decay model requires more substantial reductions in emissions than did the geometric decay model of the main text (top left). In particular, we now see that the nonnegativity constraint on net emissions binds throughout the twenty-second century. Around the year 2275 (past the end of the plot), the nonnegativity constraint stops binding. Abatement reaches $A(\tau)$ and temperature reaches $\bar{T}$ very shortly thereafter.

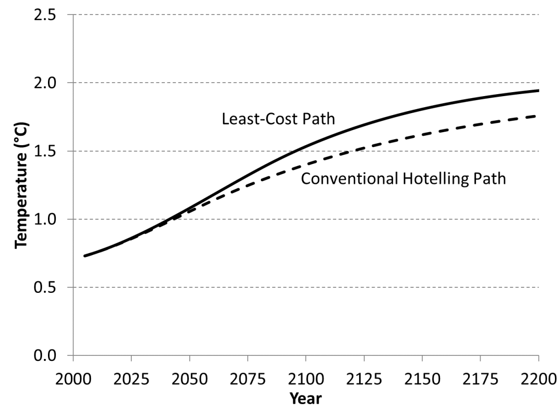
As in the main text, we see that the policymaker takes advantage of inertia to allow total CO_2 to overshoot $\bar{M}$, but now the magnitude of overshoot is reduced (bottom left). The dotted lines show that the overshoot is due entirely to the decaying stock $M_2(t)$. The non-decaying stock $M_1(t)$ cannot overshoot because it can never decline.

The bottom-right panel shows that the least-cost emission tax increases until abatement is equal to business-as-usual emissions. At this point, there are no more net emissions and abatement cannot rise further. The least-cost path increases slower than exponentially: the

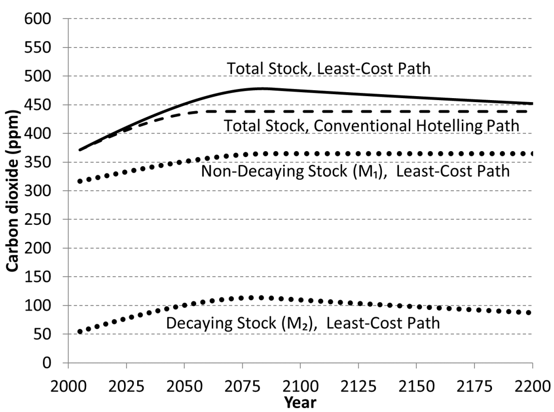
[15]The kink in emissions under the Hotelling trajectory is due to reaching $\bar{M}$.



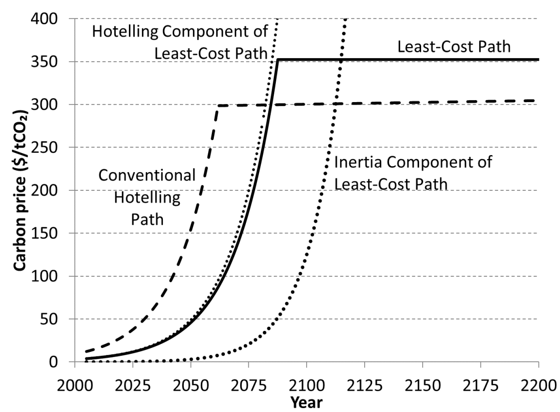
(a) Emissions net of abatement



(b) Temperature



(c) Carbon dioxide



(d) Carbon price

Figure D1: The least-cost trajectories (solid lines) for emissions, temperature, CO_2, and the carbon price for a temperature limit of $\bar{T} = 2$°C, using the carbon model of Golosov et al. (2014). Also, the conventional Hotelling-like paths (dashed lines), which are also the least-cost paths for the corresponding CO_2 constraint.

exponentially increasing Hotelling component (dotted) of the least-cost tax quickly exceeds the actual least-cost tax. Prior to reaching the maximum allowed value for abatement, the least-cost path is equal to its Hotelling component minus its inertia component. After reaching that maximum allowed value, the shadow value of the constraint accounts for the gap between the maximum allowed emission tax and the emission tax implied by the Hotelling and inertia components. We see that the inertia component becomes large around the same time that the constraint on net emissions begins to bind. This growing inertia component works to offset the exponentially increasing Hotelling component. At first, the gap between the maximum allowed emission tax and the shadow cost of emissions (which is the Hotelling component minus the inertia component) grows. This means that the shadow value of the constraint grows after it first binds. However, as the inertia component grows, that gap shrinks. The shadow value of the constraint then also begins declining, eventually falling back to zero. At that time (past the end of the plot), the constraint stops binding and temperature soon reaches $\bar{T}$.

Table D1 is the analogue of Table 1 in the main text. The new decay model restricts the policymaker much more severely than did the geometric decay model: some fraction of CO_2 now persists forever, so the policymaker must reduce emissions more aggressively to make up for the reduction in natural decay. Accordingly, all temperature limits imply a much more expensive policy than estimated in the main text. The conventional Hotelling trajectory is now around three times as expensive as the least-cost trajectory. In the main text's setting, recognizing the climate system's inertia saved a bit over $2 trillion in unnecessary costs when the temperature limit was 2 degrees Celsius. Here the savings are even greater: almost $13 trillion. The new decay setting increases the magnitude of spending and also the gains from getting policy right. As in the main text, recognizing inertia allows the policymaker to use a smaller emission tax in early years, reducing the initial emission tax to less than one-third of the Hotelling value when the temperature limit is 2 degrees Celsius. The emission tax still eventually reaches a higher level along the least-cost path than along the Hotelling path, but the percentage increase in the peak emission tax was greater in the main text's setting because there CO_2 was able to overshoot its steady-state level by a larger amount. In contrast, the presence of a permanent CO_2 stock here forces the policymaker to be less aggressive in overshooting the steady state level of CO_2.

Finally, while cumulative abatement over an infinite horizon is now fixed by the temperature limit, recognizing inertia does still allow the policymaker to reduce cumulative abatement over the next 200 years. The savings in the next 200 years' cumulative abatement are of roughly similar magnitudes as in the main text's setting, even though the required abatement is about twice as great. Therefore the savings as a percentage of cumulative abatement are much smaller here than in the main text's setting. Finally, note that because cumulative abatement over an infinite horizon is now fixed by $\bar{T}$, the monetary savings from using the least-cost path must ultimately be driven by discounting in the new decay setting.

Table D1: The present cost of each policy program, the initial carbon prices, the peak carbon prices, and cumulative abatement over the next 200 years, using the carbon model of Golosov et al. (2014).

	Temperature limit (°C)		
	2	**2.5**	**3**
Cost of efficient path from 2005–2205 ($billion)	6,750	1,046	130
Cost of Hotelling path from 2005–2205 ($billion)	19,397	3,994	686
CO_2 price along the efficient path in 2005 ($/t$CO_2$)	3.91	0.59	0.07
CO_2 price along the Hotelling path in 2005 ($/t$CO_2$)	12.24	2.35	0.39
Peak CO_2 price along the efficient path ($/t$CO_2$)	353	353	353
Peak CO_2 price along the Hotelling path ($/t$CO_2$)	305	289	271
Abatement from 2005–2205 along the efficient path (Gt C)	1,474	1,147	778
Abatement from 2005–2205 along the Hotelling path (Gt C)	1,551	1,281	986

E Nonstationary business-as-usual emissions

We now relax the assumption that business-as-usual emissions are constant. Let these emissions evolve exogenously, as $E(t)$. It is easy to see that the only necessary condition that changes in an interesting way is the condition that $h^2 = 0$, which gave us $A(\tau)$.[16] Our qualitative conclusions about the role of inertia in least-cost policy are therefore unchanged. The new condition that $h^2 = 0$ now pins down $A(t)$ for $t \geq \tau$ as

$$A(t) = E(t) - \delta(\bar{M} - M_{pre}).$$

We model the emission nonnegativity constraint as in Section D, which modifies the Maximum Principle's necessary condition to

$$C'(A(t)) = \lambda_M(t) - \mu(t) - \nu(t)\phi\, s\, F'(M(t)), \tag{E-17}$$

with $\mu(t) \geq 0$, $A(t) - E(t) \leq 0$, and $\mu(t)\,[A(t) - E(t)] = 0$.

To numerically solve this nonstationary setting, we guess τ and $\lambda_T(\tau^-)$. The guess for τ gives us $A(\tau)$ and thus $\lambda_M(\tau^-)$. We know $M(\tau) = \bar{M}$ and $T(\tau) = \bar{T}$. We solve the system

[16]Note in particular that h^0, h^1, and h^2 are unchanged except for the dependence of E on t in h^2. Even though we now have explicit time dependence in the problem, the other necessary conditions for a least-cost trajectory are unchanged because, from Hartl et al. (1995), the only partial derivatives with respect to time that would matter are those of h^0 and h^1, which are still zero.

backwards until reaching either time t_0 or a time when the nonnegativity constraint on net emissions binds. In the latter case, we then simulate the constrained system backwards until the shadow value of the constraint returns to zero (or to time t_0, whichever is later), and then simulate the unconstrained system backwards to time t_0.[17] During the interval for which the constraint binds, the shadow value of the constraint evolves according to the differential equation found by differentiating equation (E-17). Once we have reached t_0, we compare $T(t_0)$ and $M(t_0)$ to T_0 and M_0. We iterate until our guesses for τ and $\lambda_T(\tau^-)$ yield paths that satisfy the initial conditions.

We calibrate the evolution of business-as-usual emissions to total CO_2 emissions in the DICE model, with investment optimized and abatement fixed at zero. This calibration yields the following relationship for business-as-usual emissions, with emissions in Gt C and time in years:

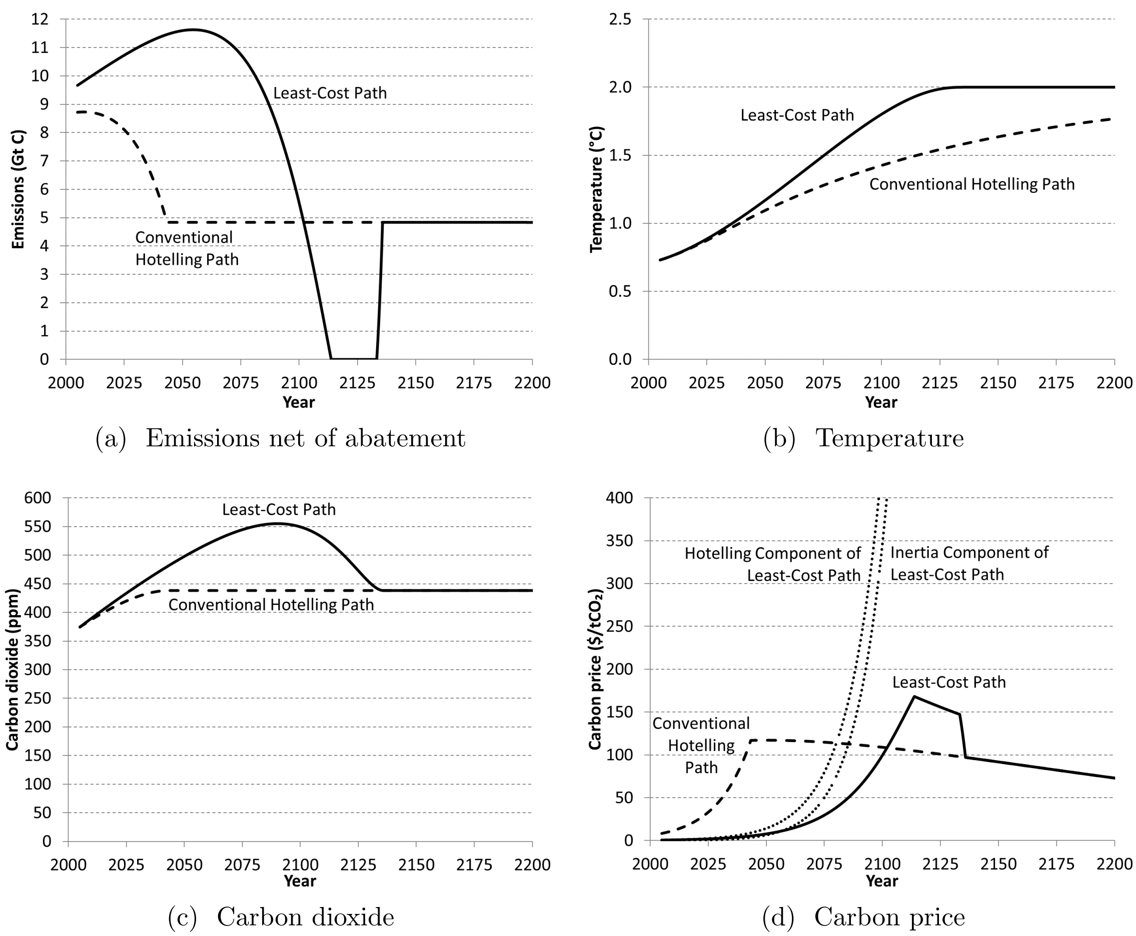
$$E(t) = 9.9662\, e^{0.0068(t-t_0)}.$$

This calibration has business-as-usual emissions increasing over time.

Figure E2 is the analogue of Figure 1 in the main text. We see that the least-cost path now has net emissions increase over the next 50 years, as business-as-usual emissions increase faster than does abatement (top left). However, abatement ramps up quickly near the end of the century, so that net emissions fall rapidly and the nonnegativity constraint begins to bind early in the next century. As in all other cases, the Hotelling policy abates emissions too aggressively over the next decades. As a result, temperature increases faster under the least-cost policy trajectory (top right). CO_2 overshoots its steady-state level $\bar{M}$ by a larger amount than in the setting with stationary emissions (bottom left). Finally, the bottom right panel shows the efficient emission tax and its components. As in Figure D1, the shadow cost of emissions along the least-cost trajectory is the difference between the Hotelling and inertia components. The efficient emission tax equals this shadow cost as long as the nonnegativity constraint on net emissions does not bind, and once that constraint binds, the shadow value of the constraint picks up the difference between the shadow cost of emissions and the maximal allowed emission tax. Once the constraint begins binding, its shadow value grows, but its shadow value eventually falls as the inertia component becomes larger relative to the Hotelling component (which makes the shadow cost of emissions fall back towards the maximal allowed emission tax). After the constraint ceases to bind, abatement quickly moves to $A(\tau)$ and temperature reaches $\bar{T}$.[18]

[17] See Section D for more on handling this constraint.

[18] Note that the efficient emission tax declines during the period in which the nonnegativity constraint binds and also in the period after τ, during which abatement holds CO_2 at $\bar{M}$. In these intervals, the change in abatement is exactly equal to the change in emissions ($\dot{A}(t) = \dot{E}(t)$). From equation (C-4), marginal abatement cost changes over these intervals in proportion to $(a_2-1)\,\dot{A}(t)/A(t) - a_2\,\dot{E}(t)/E(t)$. Thus, marginal abatement cost (but not total abatement cost) declines over these intervals if $a_2[E(t) - A(t)] - E(t) \leq 0$, which holds as long as $A(t)$ is not too much smaller than $E(t)$. Allowing Y to increase with business-as-usual emissions in equation (C-4) would introduce a force that would make marginal abatement cost more likely



(a) Emissions net of abatement

(b) Temperature

(c) Carbon dioxide

(d) Carbon price

Figure E2: The least-cost trajectories (solid lines) for emissions, temperature, CO_2, and the carbon price for a temperature limit of $\bar{T} = 2°C$, with business-as-usual emissions increasing over time. Also, the conventional Hotelling-like paths (dashed lines), which are also the least-cost paths for the corresponding CO_2 constraint.

Table E2: The present cost of each policy program, the initial carbon prices, the peak carbon prices, and cumulative abatement over the next 200 years, with business-as-usual emissions increasing over time.

	Temperature limit (°C)		
	2	**2.5**	**3**
Cost of efficient path from 2005–2205 ($billion)	604	122	26
Cost of Hotelling path from 2005–2205 ($billion)	5,489	1,475	379
CO_2 price along the efficient path in 2005 ($/t$CO_2$)	0.65	0.10	0.02
CO_2 price along the Hotelling path in 2005 ($/t$CO_2$)	8.38	1.57	0.28
Peak CO_2 price along the efficient path ($/t$CO_2$)	168	138	114
Peak CO_2 price along the Hotelling path ($/t$CO_2$)	117	88	69
Abatement from 2005–2205 along the efficient path (Gt C)	2,879	2,410	1,931
Abatement from 2005–2205 along the Hotelling path (Gt C)	3,167	2,783	2,384

Table E2 is the analogue of Table 1 in the main text. Unsurprisingly, having business-as-usual emissions increase exogenously raises the total cost of the policy program, though policy is still cheaper than in Section D where we used the carbon model of Golosov et al. (2014). The savings from using the least-cost policy are now greater than in the main text, so that ignoring inertia now costs almost $5 trillion under a 2 degree Celsius target (as opposed to just over $2 trillion in the main text). We again see that the least-cost policy uses a much lower initial carbon price and a much greater peak carbon price than does the Hotelling policy. Assuming that business-as-usual emissions increase exogenously leads to greater cumulative abatement under either policy. However, we see that using the least-cost policy instead of the Hotelling policy now reduces cumulative abatement by an even greater amount than in the main text's setting with stationary business-as-usual emissions.

F Alternate degrees of inertia and discounting

Figure F3 shows how the strength of inertia (left column) and the choice of discount rate (right column) affect the least-cost trajectories for achieving a 2°C temperature limit. Reducing inertia (i.e., increasing ϕ) means that the least-cost policy has to reduce emissions faster in order to avoid $\bar{T}$: temperature increases faster than in the baseline case even as

to be increasing over these two intervals.

CO_2 follows a lower trajectory (dashed lines). In contrast, increasing inertia (i.e., reducing ϕ) means that the effect of current CO_2 on temperature is delayed even further. The initial portion of the emission price trajectory is therefore lower and, in line with our analytic results, flatter. CO_2 now peaks over 100 ppm above $\bar{M}$ (dotted lines) even as temperature remains further from $\bar{T}$. However, even though increasing inertia lowers the initial carbon price, it does strongly raise the eventual peak carbon price (beyond the end of the plotted period) because the high degree of overshoot in CO_2 requires more aggressive abatement in order to return to $\bar{M}$.

The right column of Figure F3 shows the implications of reducing the annual consumption discount rate from the value of 5.5% used in DICE-2007 to the value of 1.4% used in Stern (2007). By raising the present cost of each unit of future abatement, the lower discount rate flattens the carbon price trajectory, which raises this century's carbon prices and lowers the next century's carbon prices. The initially higher carbon prices imply greater abatement early on, which lowers both the CO_2 and temperature trajectories. By increasing the present cost of future abatement, the lower discount rate reduces the economic importance of inertia. The more that CO_2 overshoots $\bar{M}$, the more abatement will eventually be needed to bring it back down to $\bar{M}$ before temperature reaches $\bar{T}$ (i.e., the higher the spike in the carbon price seen in the figures' bottom rows). As a result, the least-cost CO_2 trajectory overshoots $\bar{M}$ by only around 50 ppm under the lower discount rate, less than two-thirds of the overshoot under the higher discount rate, and the policy path is less peaked than with the higher discount rate.

G Phase portrait analysis

We now return to the setting and results of the main text. We construct conditional phase portraits for $t < \tau$ in order to better understand the evolution of abatement and CO_2 along a least-cost trajectory. Figure G4 depicts conditional phase portraits for a period with low temperature (top panel) and for a period with high temperature (bottom panel). These two snapshots correspond, respectively, to the early part of this century and to sometime late in this century or early in the next. The emission price (λ_M) is on the vertical axes, and CO_2 (M) is on the horizontal axes. Let $a(\cdot)$ denote the inverse of marginal abatement cost, so that $A(t) = a(\lambda_M(t))$. By the properties of $C(\cdot)$, we have that $a(0) = 0$ and $a'(\cdot) > 0$.

In each panel, the downward-sloping solid curve depicts, from equation (1), the M-nullcline:

$$M(t)|_{\dot{M}(t)=0} = \frac{1}{\delta}\left[E - a(\lambda_M(t))\right] + M_{pre}.$$

At these combinations of CO_2 and abatement, the CO_2 concentration is stationary. Decay increases in CO_2, so higher levels of CO_2 become stationary at lower levels of abatement. This curve is linear if abatement cost is quadratic. The downward-sloping dashed curve in

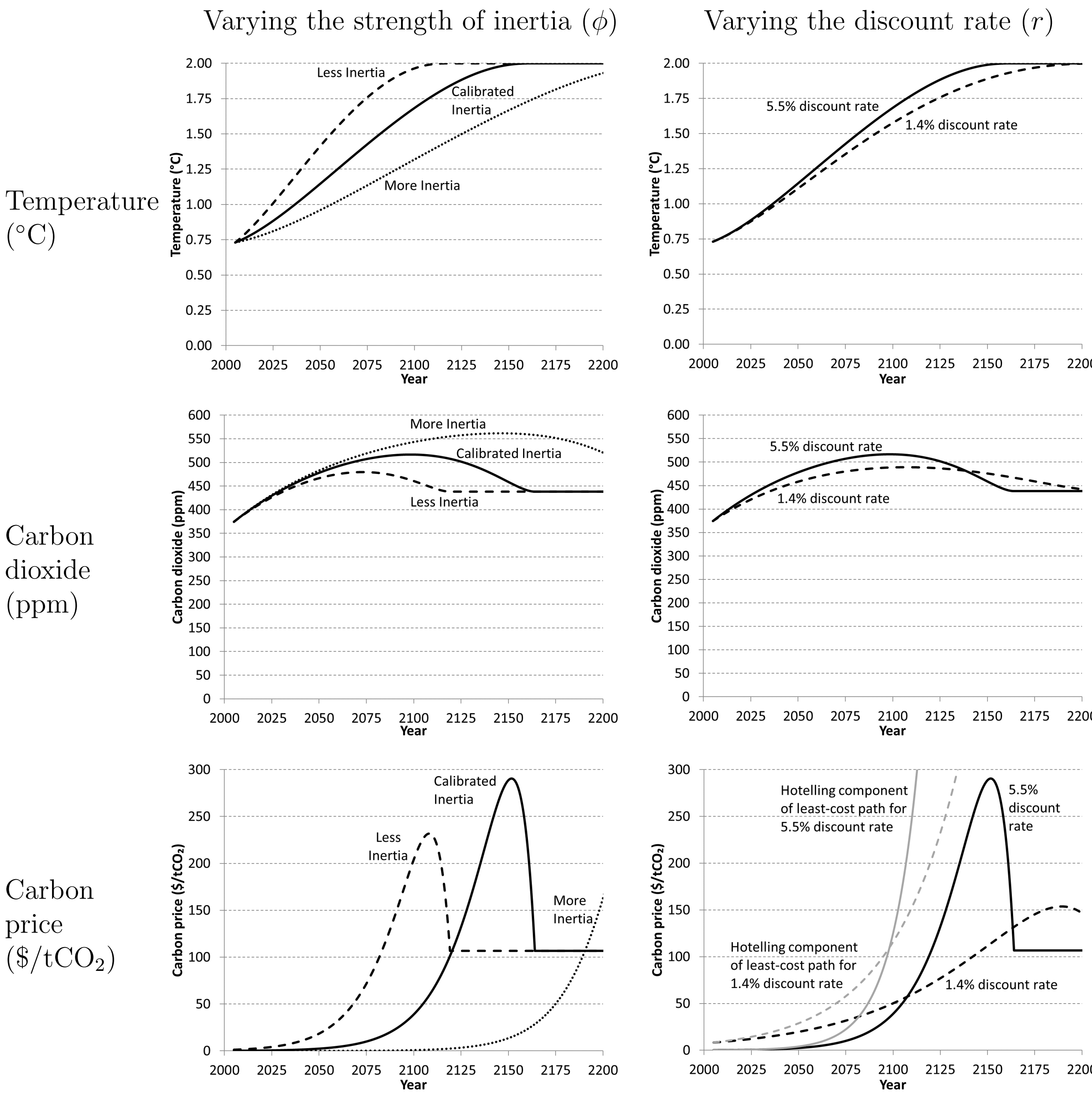
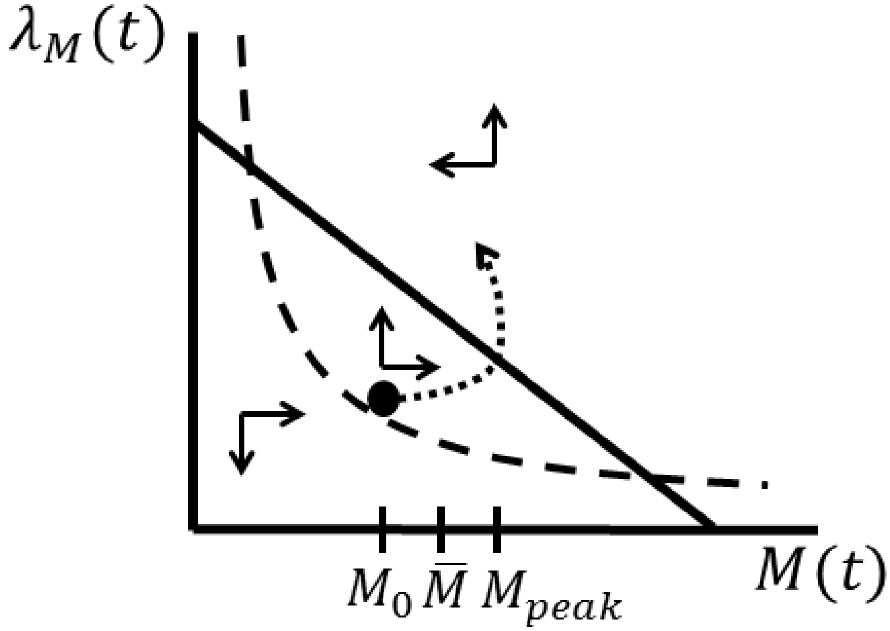
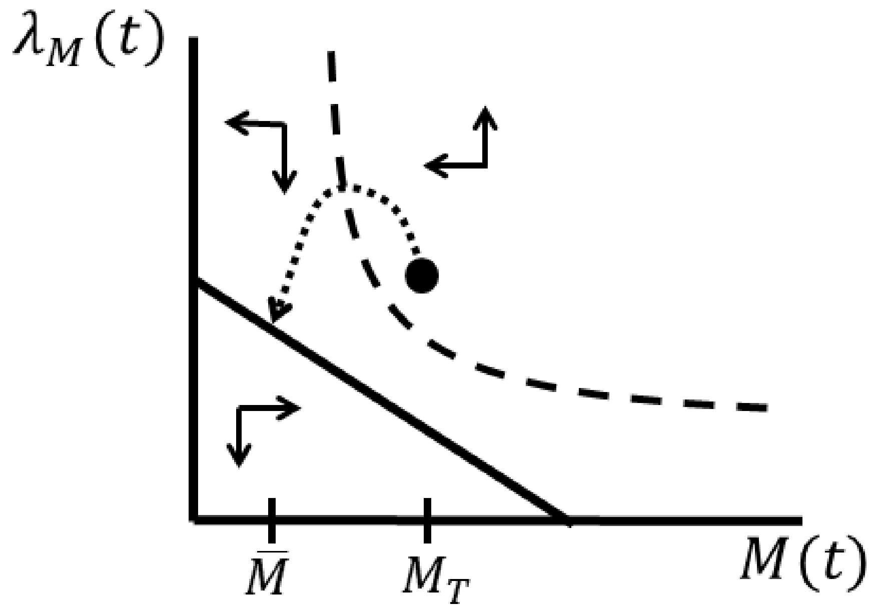


Figure F3: The least-cost trajectories for temperature, CO_2, and the carbon price for a temperature limit of $\bar{T} = 2$°C. The solid lines show the paths under the baseline calibration. In the left column, dashed lines double ϕ to 0.0182 and dotted lines halve ϕ to 0.0046 (from the baseline value of 0.0091). In the right column, dashed lines lower r to 0.014 (from the baseline value of 0.055).



(a) Near-term



(b) Long-term

Figure G4: Phase portraits conditional on λ_T and $t < \tau$. Solid curves give the M-nullclines, dashed curves give the λ_M-nullclines, dotted curves depict least-cost trajectories, and arrows give the direction of motion in each sector. The top panel corresponds to a case with $T(t)$ sufficiently far below $\bar{T}$, and the bottom panel corresponds to a case with $T(t)$ closer to $\bar{T}$.

each panel depicts, from equation (5), the λ_M-nullcline:

$$\lambda_M(t)|_{\dot{\lambda}_M(t)=0} = \frac{\phi\, s}{r+\delta} F'(M(t))\, \lambda_T(t) = e^{(r+\phi)(t-t_0)} \frac{\phi\, s}{r+\delta} F'(M(t))\, \lambda_T(t_0).$$

At these combinations of CO_2 and abatement, a least-cost trajectory holds abatement constant. The nullcline's convexity arises from using the scientific result that $F''(M(t)) < 0$, and the nullcline shifts out as the shadow cost of temperature increases. The arrows describe the direction of motion in each sector. They follow from recognizing that

$$\frac{\partial \dot{M}(t)}{\partial \lambda_M(t)} < 0 \text{ and } \frac{\partial \dot{\lambda}_M(t)}{\partial M(t)} > 0,$$

where we again use $F''(M(t)) < 0$. In sectors above (below) the M-nullcline, the direction of motion is to the left (right). In sectors to the right (left) of the λ_M-nullcline, the direction of motion is upward (downward).

The top panel depicts a case in which the nullclines intersect: business-as-usual emissions are sufficiently great that the M-nullcline is pushed out, and temperature is sufficiently far below $\bar{T}$ that its shadow cost is low and the λ_M-nullcline is pushed in. This case corresponds to the present day for a sufficiently lax temperature target. The point M_0 depicts a typical starting point, and $\bar{M} > M_0$ indicates the steady-state level of CO_2 corresponding to $\bar{T}$. The optimal emission price begins by following the dotted curve. It starts at a relatively low level in the space between the two nullclines, and it increases along with CO_2. It eventually crosses the M-nullcline at M_{peak}, at which point CO_2 begins to fall even as abatement continues increasing. This crossing illustrates how the least-cost CO_2 trajectory temporarily overshoots the terminal level $\bar{M}$.

As time passes, the shadow cost of temperature increases and the λ_M-nullcline shifts out.[19] Eventually we reach a situation such as the bottom panel, where the two nullclines no longer intersect. This corresponds to a world like that in the next century, once temperatures are closer to the chosen limit and once technological change has potentially lowered business-as-usual emissions. It also corresponds to the present world under a sufficiently stringent temperature target. In this panel, CO_2 has already peaked. The story from the last panel finished at a point such as M_T, where we pick up in this panel. As already noted, abatement is increasing and CO_2 is decreasing. The terminal condition has the policymaker hitting the M-nullcline at $\bar{M}$. As CO_2 falls, the system crosses the λ_M-nullcline. At this point, abatement peaks. As the policymaker steers the system towards $\bar{T}$, she decreases abatement towards the level compatible with steady-state $\bar{M}$.

In sum, we have seen that the type of CO_2 trajectory depends on the stringency of the temperature limit. For a sufficiently lax limit, least-cost policy increases CO_2 past its terminal level, relying on the climate system's inertia to avoid crossing $\bar{T}$. It then decreases CO_2

[19] And if business-as-usual emissions exogenously decrease, then the M-nullcline shifts in.

back towards its terminal level, using both abatement and natural decay. For a sufficiently stringent target, CO_2 begins far enough past its terminal level that abatement policy immediately begins decreasing CO_2. In either case, least-cost abatement policy generally increases before decreasing. This least-cost abatement trajectory looks quite different from the conventionally assumed, monotonically increasing Hotelling-like trajectory, and the least-cost CO_2 trajectory looks quite different from the CO_2 trajectory implied by capping concentrations at the terminal level $\bar{M}$.

Finally, consider how the least-cost CO_2 trajectory changes with properties of the climate system. In the top panel, whether CO_2 initially increases or decreases depends on how M_0 corresponds to the gap between the nullclines. For sufficiently large M_0, abatement begins at a sufficiently high level to decrease CO_2. This case is more likely the larger are ϕ, s, $F'(M(t))$, and $\lambda_T(t_0)$. For a given temperature, larger ϕ (i.e., lower inertia) increases the speed with which warming responds to any CO_2 in excess of $\bar{M}$. Larger s and $F'(M(t))$ increase the effect of CO_2 on temperature, which decreases $\bar{M}$ and so increases the degree to which M_0 is overshooting $\bar{M}$. Finally, greater $\lambda_T(t_0)$ corresponds to a more stringent temperature target, which also decreases $\bar{M}$ and increases the degree of overshoot from M_0.

H Least-cost geoengineering trajectory

The only way to achieve a CO_2 target is to reduce emissions or, perhaps, to suck CO_2 directly out of the atmosphere, but a temperature target could be achieved by directly reducing forcing. Geoengineering methods for reducing forcing typically involve "solar radiation management": if we reflect incoming solar radiation by injecting particles into the atmosphere, by placing mirrors in space, or by brightening the tops of clouds, then we can reduce forcing without reducing greenhouse gases. These methods are drawing increasing attention because they are potentially cheap but also potentially full of surprises and side-effects (Keith, 2000; Shepherd, 2012; Caldeira et al., 2013).

We here extend the theoretical setting by allowing for a geoengineering control in the form of solar radiation management. The time t level of the control is $G(t) \geq 0$, and the cost of exercising the control is a strictly increasing, convex function $D(G)$, where $D(0) = 0$. The geoengineering control reduces contemporaneous forcing, which changes the temperature transition to

$$\dot{T}(t) = \phi \left[s \left\{ F(M(t)) - G(t) \right\} - T(t) \right]. \tag{H-18}$$

The policymaker's objective is to select abatement and geoengineering trajectories in order to minimize the present cost of maintaining temperature weakly below $\bar{T}$:

$$
\begin{aligned}
V(M(t_0), T(t_0), t_0) = \min_{A(t),G(t)} \int_{t_0}^{\infty} e^{-r(t-t_0)} \left[C(A(t)) + D(G(t))\right] dt \\
\text{subject to equations (1) and (H-18),} \\
T(t) \leq & \bar{T}, \\
A(t) \geq & 0, \\
G(t) \geq & 0, \\
M(t_0) = M_0,\ & T(t_0) = T_0.
\end{aligned}
$$

The current-value Hamiltonian becomes:

$$
\begin{aligned}
H(M(t), T(t), A(t), G(t), \lambda_M(t), \lambda_T(t)) = & C(A(t)) + D(G(t)) \\
& + \lambda_M(t) \left[E - A(t) - \delta \left(M(t) - M_{pre}\right)\right] \\
& + \lambda_T(t)\, \phi \left[s \left\{F(M(t)) - G(t)\right\} - T(t)\right].
\end{aligned}
$$

The necessary conditions are unchanged, except that the new temperature transition equation must be obeyed and there is now an additional condition:

$$
D'(G(t)) = \lambda_T(t)\, \phi\, s - \nu(t)\, \phi^2 s.
$$

For times $t < \tau$, we have $\nu(t) = 0$. Therefore, for $t < \tau$, the marginal cost of geoengineering along a least-cost path increases with the shadow cost of temperature, which we have seen increases exponentially at rate $r + \phi$. Intuitively, the goengineering control directly affects temperature, so an efficient policy pathway equates its marginal cost to the shadow cost of temperature. And we have already seen that the shadow cost of temperature grows at rate $r + \phi$, reflecting both the time benefit and the inertial benefit of delaying a unit of warming.

References from the Appendix

Caldeira, Ken, Govindasamy Bala, and Long Cao (2013) “The science of geoengineering,” *Annual Review of Earth and Planetary Sciences*, Vol. 41, No. 1, pp. 231–256.

Caputo, Michael R. (2005) *Foundations of Dynamic Economic Analysis: Optimal Control Theory and Applications*: Cambridge University Press.

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski (2014) “Optimal taxes on fossil fuel in general equilibrium,” *Econometrica*, Vol. 82, No. 1, pp. 41–88.

Hartl, Richard F., Suresh P. Sethi, and Raymond G. Vickson (1995) “A survey of the maximum principles for optimal control problems with state constraints,” *SIAM Review*, Vol. 37, No. 2, pp. 181–218.

Judd, Kenneth L. (1998) *Numerical Methods in Economics*, Cambridge, Mass.: MIT Press.

Keith, David W. (2000) "Geoengineering the climate: History and prospect," *Annual Review of Energy and the Environment*, Vol. 25, pp. 245–284.

Lemoine, Derek and Christian Traeger (2014) "Watch your step: Optimal policy in a tipping climate," *American Economic Journal: Economic Policy*, Vol. 6, No. 1, pp. 137–166.

Nordhaus, William D. (2008) *A Question of Balance: Weighing the Options on Global Warming Policies*, New Haven: Yale University Press.

Ramaswamy, V., O. Boucher, J. Haigh, D. Hauglustaine, J. Haywood, G. Myhre, T. Nakajima, G.Y. Shi, and S. Solomon (2001) "Radiative forcing of climate change," in J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson eds. *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, p. 881.

Shepherd, J. G. (2012) "Geoengineering the climate: an overview and update," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 370, No. 1974, pp. 4166–4175.

Stern, Nicholas (2007) *The Economics of Climate Change: The Stern Review*, Cambridge: Cambridge University Press.