

Innovation and the Dynamics of Global Warming*

Ralph A. Winter
Sauder School of Business, UBC

May 15, 2011

This version: Feb 7 2012

Abstract

Global warming and greenhouse gases are a dynamic system with positive feedback effects. Fossil fuels are exhaustible resources. These two facts mean that innovation in clean energy technology can lead to a permanently higher temperature path. Innovation subsidies, unsupported by carbon pricing, can make global warming worse, not better. This paper explores the impact of innovation in the simplest model linking the economic theory of exhaustible resources with positive feedback dynamics in the carbon cycle.

*I am grateful to Werner Antweiler, Jim Brander, Nancy Gallini, Devon Garvie, Shreekanth Gupta, Shi-Ling Hsu, Andrew Leach Mahesh Nagarajan and numerous seminar participants for helpful comments. I am especially grateful to Don Fullerton for a thoughtful reading of an earlier draft. I thank Kairong Xiao, Zhe Yuan and Zin Zhao for excellent research assistance and the Social Sciences and Humanities Council of Canada for research support.

1 Introduction

The innovation and development of clean energy sources such as wind and solar energy are emerging as a key strategy in the battle against global warming. The strategy rests on a seemingly obvious proposition: innovation lowers the cost of clean energy, leading to substitution away from fossil fuels, which lowers carbon emissions and mitigates the problem of global warming.

The proposition, unfortunately, is false. Innovation in clean energy can set global temperatures on a permanently higher path. The subsidy of innovation, as a naked policy instrument unsupported by carbon pricing, is not merely suboptimal policy. Subsidizing innovation can make global warming worse.

To develop the economic relationship between clean-energy innovation and climate change, I start with a paradox familiar to environmental economists. Fossil fuels are an exhaustible resource. Suppose that tomorrow a clean, inexhaustible energy substitute were universally available at a cost equivalent to 60 dollars per barrel of oil. The owner of any conventional fuel deposit with low extraction costs would prefer to sell at 59.99 or less rather than share the energy market with the substitute. Oil from these deposits will therefore be sold before clean energy captures any market share and at lower prices as a result of the innovation. The effect of the innovation in clean energy is that fuel will be exhausted – and carbon emitted – more intensively and at an earlier date. This paradox is that carbon emissions are initially higher as a result of clean energy innovation.¹

As set out in the literature, however, the theory predicts that in the long run clean energy innovation helps in the battle against global warming. Innovation in clean energy has two effects on carbon emissions. First, as in the example above, carbon is released earlier into the atmosphere as a result of innovation. In existing models, this early release of carbon into the atmosphere is neutral or beneficial in terms of the long run temperature trend as the atmospheric carbon is reabsorbed into the earth’s surface at a steady rate. The environment, represented by a single state variable, is able to improve over time. The second effect of innovation is that fossil

¹The “green paradox” literature in global warming develops two related themes. The first is that rising taxes or a commitment to high future taxes will have the perverse effect of raising current emissions (Sinclair (1992), Sinn (2008), Hoel (2012)). This type of perverse initial reaction to the announcement of future tax changes is common to any economic activity with inter-temporal substitution. (An announced future tax on housing will increase current housing activity, for example.) The second theme is the perverse effect of innovation (Strand (2007), Hoel (2008)). Hoel, for example, shows that a reduction in the costs of clean energy will initially worsen global warming as current emissions rise.

fuel deposits with high extraction cost (above 60 dollars in our example) will be left in the ground rather than extracted. Less carbon is emitted into the atmosphere – a clear benefit of clean-energy innovation. The net effect is that the paradox disappears in the long run. The prediction is that innovation eventually works as intended.²

This prediction is too optimistic. The theory offered here represents carbon in the biosphere via *two* state variables, carbon in the atmosphere and carbon on the earth’s surface. This allows us to include a fundamental feature of carbon cycle dynamics: positive feedback effects. As greater atmospheric carbon raises the global temperature, reflective ice-fields melt and methane gas is released from melting permafrost (to take just two examples), resulting in a higher *rate* of flow of carbon to the atmosphere. The effect is that an initial increase in carbon emissions that raises global temperature will increase the rate at which carbon escapes from the earth’s surface and accumulates in the atmosphere. Innovation combined with the sufficiently strong feedback effects then yields higher temperature paths not just in the short run but permanently. The acceleration of carbon emissions (the first effect of innovation) may overwhelm even in the long run the benefit of reduced total carbon emissions (the second effect). Because of positive feedbacks, even a small innovation may lead the temperature path to a discretely higher steady-state temperature. Global warming is a long run problem and it is the long run consequences of global warming policies that are critical for policy.

The theory here argues against clean energy innovation subsidies as a naked policy instrument. As a component of a portfolio of policies, however, clean-energy innovation subsidies are of value because the other main policy instrument, carbon pricing, eliminates the “dark side” of innovation. To render innovation of value, carbon prices must be *reactive* to innovation successes – and reactive in a non-obvious way. When a new innovation gives clean energy producers an advantage over conventional energy, a reactive carbon price policy is often one that magnifies this advantage, by raising the tax on fossil fuel use. Fossil fuel producers are hit with a double whammy.

Carbon pricing is thus an important *complementary* instrument to clean energy innovation, being necessary even to ensure that the net impact of innovation is positive. This complementarity is not well understood among policymakers. With carbon taxes seemingly impossible to implement given U.S. politics, clean-energy research and development is becoming the focus across a range of the political spectrum.³ Support for the policy trend presupposes that carbon

²In Hoel (2008) the long-run impact of clean energy innovation is to lower temperature whenever fossil fuel deposits with high extraction costs are displaced by clean energy.

³Consider, for example, the recent joint call by the Brookings Institute and the American Enterprise Institute

pricing and clean energy subsidies are *substitutes* in the battle against global warming. This is natural assumption, given that these are two instruments available to solve the same problem, but the assumption is misguided. Carbon pricing is even more essential when clean energy innovation is successful than when it is not.

This paper contributes to a growing literature on global warming and an earlier literature on exhaustible resource economics. The clean energy paradox, or green paradox, in terms of the impact of innovation described is developed in Strand (2007) and very clearly in Hoel (2008) as discussed in footnote 1. Acemoglu et al (2011) develop a dynamic model integrating the economy with climate change in which the driver of policy design is endogenous technical change. These authors assume a constant rate of environmental regeneration. This assumption would allow the environment to recover completely from any past damages if the rate of emissions could be reduced sufficiently. It is never too late to recover. The built-in optimism about the long run in Acemoglu et al (2011) and in Hoel (2008) disappears once the positive-feedback mechanisms of the carbon cycle are recognized. Positive feedback mechanisms require that the environment be represented by multiple state variables, as in the climate science literature, not a single state variable as in Acemoglu et al and Hoel.⁴

I allow uncertainty in the innovation process in my model to capture an important distinction between the ex ante effect on global warming of the *threat* of innovation and the ex post effect of successful innovation. In doing so, I draw on the economic literature on the theory of pricing in an exhaustible resource market subject to innovation in a backstop technology (Heal 1976, Gallini, Lewis and Ware 1983, Dasgupta and Stiglitz 1981) as well as the literature on exhaustible resource pricing with heterogeneous extraction costs or reserve-dependent costs (see Devarajan and Fisher 1981).

The possibility of perverse effects of clean-energy innovation and the complementarity of innovation and carbon pricing are not fully appreciated in the climate policy literature. The

for an increase in clean energy investment from 4 billion to 25 billion annually (Hayward et al, 2010). In the popular press as well, clean energy subsidies have been touted as a superior instrument to carbon taxes and popular support is in favour of subsidies (e.g, David Leonhardt, "A Climate Proposal Beyond Cap and Trade", New York Times, October 12, 2010).

⁴Chakravorty et al (2011), Leach (2009) and Laurent-Lucehettii and Leach (2012) are additional important contributions to the literature on innovation and climate change. A related literature explores the impact of taxes on global warming, Sinclair (1992) showing that high future taxes (rising taxes) can increase GHG emissions. (See also Sinclair (1994), Ulph and Ulph (1994), Sinn (2008).) Sinn labelled this effect the "green paradox". This term has come to refer to the paradoxical effect of either future taxes or current innovation in raising current emission levels.

most prominent model integrating economics and climate dynamics is the Dynamic Integrated Model of Climate and the Economy (the DICE model) (Nordhaus, 2008) and its extension to a regionally disaggregated model (Nordhaus 2010 and 2011). The latest version of the DICE model includes a backstop clean-energy technology, but assumes (1) a current Hotelling rent of \$0.07 per ton of carbon for fossil fuels (Nordhaus, 2007, p.31), which implies a Hotelling rent of only \$0.01 per barrel of crude oil;⁵ and (2) an extraction cost of zero (Nordhaus 2008, p.43). The effects analyzed here, in contrast, rely entirely on a *large* Hotelling rent. The DICE assumptions (1) and (2) are at odds with the observed price of 80-100 dollars per barrel.⁶ The interaction of innovation and carbon pricing policy is outside the focus of the DICE model.⁷ The Stern review of global warming policy (Stern, 2006) notes that the price reaction of the fossil fuel market may dampen the effects of policy of clean energy subsidies, but does not recognize the possibility of a negative net impact of innovation. Stern is solidly of the conventional view that the net effect of innovation in clean energy must be positive. Stern surveys the important positive feedback mechanisms in carbon models but does not connect feedback mechanisms to the reaction of fossil fuel prices to innovation, which is the analytical focus of this paper.

Section 2 of this paper offers an economic model of innovation and the dynamics of global warming. Section 3 moves from positive economic theory to normative analysis, and from the consequences of existing global warming policy to optimal policy in a first-best world. I analyze

⁵A barrel of oil weighs about 135 kg, of which 83% to 87% is carbon, meaning that a barrel has about $.85 \cdot 135 \cdot 0.001 = 0.115$ tonnes of carbon, which is $1.1 \cdot 0.115 = 0.126$ tons. A rent of \$0.07 per ton of carbon is $0.07 \cdot 0.126 =$ about 1 cent per barrel.

⁶Nordhaus' DICE model aggregates coal and oil as fossil fuels. The low Hotelling rent follows from the model's assumption that there are 900 years' of consumption of fossil fuels available at current consumption rates (Nordhaus (2007): 15). Oil and Natural gas together account for a higher rate of carbon emissions than coal: Raupach et al (2007) report that over the 2001-2004 period, solid fuels (coal), liquid fuels and gas fuels accounted for an estimated 35, 36 and 20 percent of CO₂ emissions respectively. The Hotelling dynamics of oil and gas are therefore potentially important.

⁷Nordhaus offers a critique of the subsidy approach to climate change policy that is completely different than the critique developed here:

“Because of the political unpopularity of taxes, it is tempting to use subsidies for “clean” or “green” activities as a substitute for raising the price of carbon emissions. This is an economic and environmental snare to be avoided. *The fundamental problem is that there are too many clean activities to subsidize.* Virtually everything from market bicycles to nonmarket walking has a low carbon intensity relative to driving. There are simply insufficient resources to subsidize all activities that are low emitters.” (Nordhaus 2008: 21, emphasis added).

an optimal carbon tax in the model with exogenous innovation. The first-best optimum can be implemented in a competitive market with a carbon tax, a tax that will increase in the event of early success in innovation. Optimal carbon pricing does allow clean-energy innovation to have an unambiguously positive impact on social welfare. Even in the first-best optimum, however, successful innovation leads to a worsening of global warming. The concluding section offers a critical discussion of the policy literature in light of the theory developed here.

2 The Model

2.1 Assumptions

2.1.1 Physics Assumptions

The key assumption in this paper is that the carbon-cycle is a process with positive feedback. Among the feedback mechanisms identified by climate scientists, seven are particularly important. The melting of ice sheets reduces the cooling effect that the ice has in reflecting radiation away from the Earth. This means that higher temperatures lead to an increase in the rate of change of temperature, a positive-feedback process known as the ice-albedo effect. A second mechanism is that global warming could cause the death of vegetation in regions such as the Amazonian rainforests through reduced rainfall, leading to the release of CO₂ to the atmosphere and to reductions in the absorption of CO₂ by plants. Peter Cox et al (2000) uncover a third positive-feedback mechanism: global warming can result in increased respiration from bacteria in the soil, releasing additional CO₂. The fourth positive-feedback mechanism is the release of GHG's (mainly methane) from the tundra in the arctic, mainly Eastern Siberia. A well-known study in 2007, led by University of Alaska's International Arctic Research Centre and the Russian Academy of Sciences, demonstrated the strong potential for this mechanism.⁸ The fifth mechanism is the release of methane from the oceans, in the form of methyl hydrates, potentially leading to or adding to a "runaway methane global warming". The sixth positive feedback is from the evaporation of water and the accumulation of water vapour in the upper atmosphere. Water vapour is accumulated in greater amounts in a warmer atmosphere, as we know from basic physics. This leads to a higher rate of temperature increase, water vapour (not CO₂) being the most powerful greenhouse gas. The seventh positive feedback is the reduced

⁸As stated in a the National Science Foundation press release at the time of this study, "Release of even a fraction of the methane stored in the shelf could trigger abrupt climate warming." Since the study in 2007, some estimates of the amount of trapped methane have more than doubled.

ability of oceans to absorb carbon dioxide as temperature rises. The amount of carbon in the oceans is huge and a significant disruption in rate of absorption of carbon by the oceans could overwhelm other aspects of carbon-cycle dynamics.⁹ The potential power of positive feedback mechanisms is enormous.¹⁰ Obviously, the uncertainty involved in these estimates is huge, especially given the property of positive feedback mechanisms to magnify any uncertainty in dynamic systems.

To capture positive feedback effects in the simplest way, I represent the environment with two state variables: g_t^1 , the total volume of GHG's in the atmosphere and g_t^2 , the total carbon and other GHG-potential elements (which we shall refer to simply as "carbon") in the earth's surface, including the oceans. Both of these variables are as measured in the same units, e.g., CO₂-equivalents in terms of their GHG impact. Temperature or heat, h , is a function $h = H(g^1)$ of atmospheric GHG's.¹¹ The environmental state variables are linked by the two relationships: greenhouse gases g_t^1 are absorbed (transformed to g_t^2) at a constant rate a . This recapture of atmospheric carbon is through photosynthesis net of carbon released from dead plants and absorption in the ocean. (Antweiler (2010) adopts an estimate of the absorption rate of about 2 percent per year; Uzawa (2003) assumes a higher but, Antweiler argues, unreasonable rate.) GHG's are released from the earth's surface at a *rate* that is an increasing function of temperature and therefore of the current *level* of atmospheric GHG's: $F(g_1)$. This is the positive feedback effect. The strength of positive feedback mechanisms is small at low global temperatures but, over some region, increasing in temperature. I capture this below in assumptions on the shape of F . Figure 1 summarizes the assumptions on the carbon cycle.

2.1.2 Economic Assumptions

I consider a market for energy that can be supplied by fossil fuels, an exhaustible resource, or by an existing backstop technology at a known initial cost.¹² Let x_t be the flow of resource

⁹Other positive feedback mechanisms in the carbon cycle include mechanisms on the demand side of the energy market: higher temperatures lead to greater demand for air-conditioning, greater fuel use and therefore a higher rate of carbon emissions.

¹⁰Stern (2006: p.3) summarizes the scientific evidence on these effects as indicating that they would likely amplify warming by 1-2 degrees Celcius by 2100, against an estimate of an increase of at least 3 degrees celcius if GHG emissions remained at current levels, which is within the range of predictions of the IPCC (2007).

¹¹I thus ignore the long lags between atmospheric carbon build-up and global temperatures.

¹²See Dasgupta and Stiglitz (1981) for a similar model in the special case of zero extraction costs, Heterogeneity in extraction costs, assumed here, is necessary to capture the essential tradeoff in the impact of innovation on

extraction at time t , $q(p)$ be the (stationary) demand for energy, p_t be the price of energy and b_0 be the cost per unit of producing energy with the existing backstop technology. The stock of resource extracted up to time t is $s_t = \int_0^t x_\tau d\tau$. The resource and the backstop technology are each supplied competitively, in an economy with a constant rate of interest, r . I allow for heterogeneity in the costs of extraction across resource pools, with $M(c)$ being the amount of resource that can be extracted at cost c or lower. Given the well-established result that lower-cost deposits will be extracted first, I represent the heterogeneity with an increasing function $c(s)$ that gives the marginal cost of extraction when s has been extracted to date.¹³

Innovation in new energy technology is represented by the probability of discovery of a new backstop technology that produces energy at cost $b < b_0$. Only one possible new technology exists and once it is discovered, no future innovation takes place. The heterogeneity in costs includes a positive measure of deposits with extraction costs between b and b_0 : $M(b_0) - M(b) > 0$. As in previous continuous-time models of innovation (see Reinganum (1989)) the probability of discovery of the new technology in any small interval of time, dt , is ρdt . The probability ρ is exogenous to the model and can be interpreted as an instrument of government policy. I develop and compare the equilibrium path of p_t and x_t under three scenarios: innovation is not possible (the cost of the backstop remaining forever at b_0); innovation is possible but not realized prior to exhaustion of the fossil fuel; and innovation occurs at some date τ .¹⁴

The market equilibrium is linked to the global warming dynamics by an assumption that extraction x_t results in a flow x_t of GHG's to the atmosphere. (That is, I measure fossil fuels in units of carbon emitted.) The model, in sum, is Markov with five state variables: s, g^1, g^2, h ; and a binary variable $I \in \{0, 1\}$ indicating whether or not the new innovation has been discovered.

The five-state-variable model is tractable, and in fact will be reduced to a single differential equation describing the evolution of the environment in a post-extraction phase. Two natural separabilities allow this tractability. The first is a separability between the market state variables, (s_t, I_t) and the carbon/climate variables (g_t^1, g_t^2, h_t) . The market equilibrium is not affected by

global warming.

¹³I interpret $c(s)$ as including the cost of discovery of new resource pools. I set aside the uncertainty in the outcome of exploration.

¹⁴I model the connections between markets and carbon emissions; and between carbon emissions and temperature; but I do not need to refer to the relationship between temperature and damages. Incorporating damages would only magnify the costs of an early release of carbon, because of discounting. I also set aside evolution of technology in other dimensions, such as carbon emission mitigation. Allowing for an exogenous trend in improvements in this technology would add to the cost of early release of carbon.

the evolution of (g_t^1, g_t^2, h_t) because of an assumption that global warming is a pure externality (i.e., the private cost is 0). We can therefore develop the market equilibrium completely without reference to the environment. And we can develop the global warming dynamic system taking the evolution of s as exogenous, i.e. unaffected by the evolution of the other state variables in this system. The second natural separability is over time. At date T , when extraction is finished, the entire system can be described in terms of only two state variables g^1 and g^2 . The fossil fuel market simply disappears, taking with it two state variables, x_t , and I_t . The market leaves the environment endowed with (g_T^1, g_T^2) , with $C_T = g_T^1 + g_T^2$. In some cases innovation in clean energy will reduce the total amount of terminal carbon C_T in the environment but increase the amount of carbon g_T in the atmosphere. The dynamic system may be tipped into an eventual resting point where the concentration of atmospheric carbon, and the earth's temperature, are permanently higher.

2.2 Market Equilibrium

2.2.1 Equilibrium with no innovation

The case of no innovation serves as a benchmark. Hotelling (1931) offered a now-famous arbitrage condition for competitive prices in a market for exhaustible resources. In the case of a constant extraction cost c , this condition becomes $\partial(p_t - c)/\partial t / (p_t - c) = r$, i.e. rent rises at the rate of interest. Only under this condition will a competitive supply be positive at all dates until exhaustion of the resource.

In a setting of heterogeneous extraction costs, the competitive equilibrium involves the extraction of resources pools in order of extraction costs, from lowest to highest cost pools. The arbitrage condition is replaced by the first-order condition giving the optimal date of extraction of a small unit of resource of each particular quality. Let $c'(\tilde{s})$ represent $\partial c(s)/\partial s$ evaluated at \tilde{s} , and let a dot indicate a time derivative. The first-order condition, familiar from the literature on natural resources, is the following (with the right hand side expressed in several ways):¹⁵

$$\frac{\partial(p_t - c(s_t))/\partial t}{p_t - c(s_t)} = r - \frac{\partial c(s_t)/\partial t}{p_t - c(s_t)} = r - \frac{c'(s_t)\dot{s}_t}{p_t - c(s_t)} = r - \frac{c'(s_t) \cdot q(p_t)}{p_t - c(s_t)} \quad (1)$$

Consider the opportunity cost of delaying extraction by a small interval, dt , for a competitive firm holding an infinitesimal unit of the resource of given extraction cost. If the firm extracting at t were to hold its unit for an instant dt , it would have lower extraction costs than the firms

¹⁵To prove (1), note that competitive equilibrium requires that from delaying extraction from time t to time $t + dt$ must approach 0 as $dt \rightarrow 0$. This implies

extracting at time $t + dt$. The price path must make the firm indifferent (to a first-order approximation) to waiting until $t + dt$. The rate of price increase is therefore reduced by rate of decrease of extraction costs, the second term after the first equality in (1).

Extraction of fossil fuels will stop at the date T where the cost of extraction equals the backstop technology cost. We denote with a superscript m the equilibrium variables in the case where innovation is impossible. In sum, the equilibrium with no innovation is $[p_t^m, x_t^m; T^m]$ that solves, over $[0, T^m]$: (1); $x_t^m = q(p_t^m)$; $s_{T^m}^m = \int_0^{T^m} x_\tau^m d\tau$; with boundary condition $p_{T^m}^m = c(s_{T^m}^m) = b_0$; and $x_t^m = 0$ for $t > T^m$.

2.2.2 Equilibrium with innovation

When the probability of innovation, ρ , is positive, the equilibrium price path p_t is stochastic. The sample paths that we characterize are the pre-innovation paths p_t^n and x_t^n for price and quantity, with extraction terminated at date T^n if there is no discovery; and for each possible discovery date τ of the new backstop technology, the paths $p_t^d(\tau)$ and $x_t^d(\tau)$ between τ and the exhaustion date $T^d(\tau)$ contingent upon discovery at date τ . We denote by s_t^n and $s_t^d(\tau)$, for $t \geq \tau$, the cumulative resource extracted along the pre-innovation and post-innovation paths.

The date of discovery is the only random variable in this stochastic system of equilibrium paths. The equilibrium paths after discovery at τ , $p_t^d(\tau)$ and $x_t^d(\tau)$, are identical to the model above without innovation except, of course, that the terminal price is lower. The boundary condition becomes $p_{T^d(\tau)}^d(\tau) = c(s_{T^d(\tau)}^d) = b$.

The price path p_t^n is the equilibrium price path until either discovery or the cumulative resource extracted reaches the level $s = c^{-1}(b_0)$, whichever comes first. To derive the differential equation characterizing p_t^n , note the following. At time t , competitive equilibrium requires that the future value at time $t + dt$, of the rent from extracting at time t must equal the opportunity

$$\begin{aligned}
(1 + rdt) [p_t - c(s_t)] &= p_{t+dt} - c(s_t) \\
(1 + rdt) [p_t - c(s_t)] &= [p_{t+dt} - c(s_{t+dt})] + [c(s_{t+dt}) - c(s_t)] \\
rdt [p_t - c(s_t)] &= [p_{t+dt} - c(s_{t+dt})] - [p_t - c(s_t)] + [c(s_{t+dt}) - c(s_t)] \\
rdt [p_t - c(s_t)] &= \frac{\partial (p_t - c(s_t))}{\partial t} dt + \frac{\partial c(s_t)}{\partial t} dt \\
r &= \frac{\partial (p_t - c(s_t)) / \partial t}{p_t - c(s_t)} + \frac{\partial c(s_t) / \partial t}{p_t - c(s_t)}
\end{aligned}$$

and finally

$$\frac{\partial (p_t - c(s_t)) / \partial t}{p_t - c(s_t)} = r - \frac{\partial c(s_t) / \partial t}{p_t - c(s_t)} = r - \frac{c'(s_t) \cdot q(p_t)}{p_t - c(s_t)}$$

cost of waiting until $t + dt$, for vanishingly small $dt > 0$. Under risk-neutrality, this condition leads to the following (see Appendix A):

$$r[p_t^n - c(s_t^n)] = \rho[p_t^d(t) - c(s_t^d(t)) - (p_t^n - c(s_t^n))] + \frac{\partial(p_t^n - c(s_t^n))}{\partial t} + c'(s_t^n)\dot{s}_t^n(t) \quad (2)$$

As shown in the appendix, taking the limit of this expression as $dt \rightarrow 0$ yields the condition

$$\frac{\frac{\partial}{\partial t}(p_t^n - c(s_t^n))}{p_t^n - c(s_t^n)} = r - \frac{c'(s_t^n) \cdot \dot{s}_t^n(t)}{p_t^n - c(s_t^n)} + \rho \left(\frac{p_t^n - p_t^d(t)}{p_t^n - c(s_t^n)} \right) \quad (3)$$

The rate of change of rents along the pre-innovation path must equal the interest rate minus the rate of cost savings from delay, plus the expected instantaneous rate of capital loss on deposits. T^n is given as the solution to $\int_0^{T^n} q(p_t^n)dt = s_{T^n}^n = c^{-1}(b_0)$.

I have characterized the equilibrium paths in three cases: where innovation is impossible; where innovation is possible but does not occur before the cost of extraction meets the cost of the existing backstop technology; and where innovation occurs at a realized date τ . Because the system is Markov, the equilibrium values of the endogenous variables can be expressed as functions of the state variables s and I : we let $p^m(s)$, $p^n(s)$, $p^d(s)$ be the prices as functions of s conditional upon (respectively) innovation impossible; innovation possible but not yet realized, and innovation realized.

The comparison of the price functions, in s , is straightforward, given our characterizations. From (1) and (3) and $b < b_0$, it follows directly that over the extraction phase of the market $p^d(s) < p^n(s) < p^m(s)$. It then follows from $q'(p) < 0$ that, at the same value of the state variable s , the rates of change of s (the extraction rates) can be ordered under the three equilibrium paths: $\dot{s}^m < \dot{s}^n < \dot{s}^d$. Now, at each of the respective extraction-termination dates, $T = T^m$, T^n and $T^d(\tau)$, the state variable s_T takes the values $c^{-1}(b_0)$, $c^{-1}(b_0)$ and $c^{-1}(b)$. Here $c^{-1}(b) < c^{-1}(b_0)$ since $b < b_0$. The ordering of the rates of change of s then allows us to order the termination dates. In sum,

Proposition 1 *When the current benchmark technology has a cost b_0 and innovation of a lower cost is governed by a stationary probability pdt of discovering a technology of lower cost b , the equilibrium extraction path has two phases: prior to discovery, the cumulative extraction path lies above the equilibrium path for the case where innovation is impossible. If discovery is possible but never realized, the same amount of resource, $c^{-1}(b_0)$, is extracted as if innovation were impossible, but earlier and at a more intensive rate. If discovery is made at date $\tau < T^n$, extraction is at a higher rate and terminates even earlier, but the total resource extracted (and carbon emitted) is less: $\max(s_\tau^n, c^{-1}(b)) < c^{-1}(b_0)$.*

Figure 2 depicts the equilibrium price paths for a particular realization of τ , the one random variable in this stochastic system. Note that the price path p_t^n crosses the price path p_t^d , even though the prices as functions of the state variable s are ordered. Figure 3 depicts the equilibrium extraction paths, which do not cross.

When we take this pattern of extraction to the next subsection of the paper on global warming dynamics, the possibility of innovation in clean energy technology carries a cost and a benefit in terms of its impact on global warming. If innovation is possible and *if* it actually occurs at date τ , then carbon in the amount $c^{-1}(b_0) - \max[S_\tau^n, c^{-1}(b)]$ is left in the ground rather than introduced into the environment. This is a benefit. But the carbon that is released is released earlier; temperatures always rise immediately upon discovery of the innovation. Because of the positive-feedback nature of global warming the early release may set temperature on a permanently higher path in spite of the reduction in total carbon released into the atmosphere.

If $c(s_{T^m}^m) < b$, i.e. if the marginal extraction cost is everywhere less than even the cost of energy under the new technology (for example, where marginal extraction cost is constant at $c < b$), only the negative side of the tradeoff is present. The impact of the innovation is then to leave the same amount of fossil fuel extracted, and the same amount of carbon released, but at an earlier date and with greater intensity:

Proposition 2 *If $c(s_{T^m}^m) < b$ then successful innovation leads to an earlier extraction of the same amount of the resource.*

2.3 Global-warming dynamics

2.3.1 The carbon cycle, incorporating \dot{s}_t as exogenous:

The extraction of fossil fuels provides an exogenous (to these dynamics) injection of carbon into the atmosphere at a rate \dot{s}_t . The atmospheric carbon is absorbed by the earth's surface at a constant rate a , but the rate of release of carbon from the surface to the atmosphere depends on the current temperature, through the function $m(h_t)$.

The system characterizing the evolution of the four continuous state variables (s_t, g_t^1, g_t^2, h)

is given by the following:¹⁶

$$\dot{g}_t^1 = \dot{s}_t + m(h_t)(g_t^2) - ag_t^1 \quad (4)$$

$$h_t = H(g_t^1)$$

$$\dot{g}_t^2 = ag_t^1 - m(h_t)(g_t^2)$$

Defining $F(g_t^1) \equiv m(H(g_t^1))$ allows us to reduce this system to the following couplet of dynamic equations:

$$\dot{g}_t^1 = \dot{s}_t + F(g_t^1)g_t^2 - ag_t^1 \quad (5)$$

$$\dot{g}_t^2 = ag_t^1 - F(g_t^1)g_t^2$$

The positive feedback is captured in the function $F(\cdot)$. The magnitude of this function and the size of the stock g_t^2 that is vulnerable to release to the atmosphere are key in determining whether earlier release of a given stock of carbon leads to permanently higher temperatures. “Earlier” and “higher” refer to partial orders over continuous extraction paths and temperature paths. For a given amount of fossil fuel s , and two different extraction paths of s , x^1 and x^2 , with $\int_0^\infty x_t^i dt = s$, $i = 1, 2$, we say that x^1 is an *earlier* extraction path than x^2 if $\int_0^{\hat{t}} x_t^1 dt \geq \int_0^{\hat{t}} x_t^2 dt$ for all \hat{t} , with the inequality being strict for some set of \hat{t} . (This is analogous to first-order stochastic dominance.) And define a temperature path h^1 to be *higher* than a path h^2 if $h_t^1 \geq h_t^2$ with the inequality strict on a subset of dates of positive measure.

It is clear that earlier emissions of a given stock of carbon does not always lead to a permanently higher temperature path. Suppose, for example, that the positive feedback mechanism is very weak, relative to a , and compare the release of 99 percent of a stock of carbon in year 1 with the release of 99 percent of the stock in year 100. In the year 101 the temperature will

¹⁶Obviously, this set of assumptions abstracts completely from a host of economy-climate interactions and the resulting model is by design illustrative rather than realistic. (Nordhaus’ model, in contrast, has 16 dynamic equations and 24 dynamic variables; our model will reduce climate dynamics to a single differential equation.) More state variables would yield a more realistic description or computation. These would include a deep ocean carbon reservoir; the sensitivity of surface carbon to release (dependent on the history of atmospheric temperature, not just its current value); regionally variable temperature changes (the areas near the poles are forecast to experience a much higher rate of temperature increase), and so on. Energy exchanges across adjacent reservoirs, not just carbon exchanges, are incorporated in more complex climate models. Our assumption that the release of carbon into the atmosphere is proportional to surface carbon, rather than a more realistic general increasing function; the release of water vapour, for example, is not proportional to the existing stock of water on the surface. And a more realistic theory would treat separately two types of feedback effects: higher temperature causing release of additional GHG from the earth’s surface; and the reduction in the rate at which solar energy is reflected out of the atmosphere by ice-fields.

be higher under the latter scenario because the carbon from the atmosphere will not have had time to resettle (via ag^1) into the earth's surface. The temperature paths cannot be ordered. In fact, without positive feedback, it is easy to generate examples where early release of carbon is beneficial for the medium and long run health of the environment because the environment is able to recover through the recapture of atmospheric gases through the term ag_t^1 .

I capture the assumption of a strong positive feedback in the following ($C_t = g_t^1 + g_t^2$ represents total carbon in the biosphere).

$$\text{(positive feedback)} \quad \frac{\partial}{\partial g_t^1} \{F(g_t^1)(C_t - g_t^1) - ag_t^1\} > 0 \quad (6)$$

The positive feedback condition means that if a unit of carbon is transferred from the earth's surface to the atmosphere, raising the atmospheric temperature, the *rate* of carbon release from the surface to the atmosphere increases. All of the physical processes outlined in section 2.1.1 combine to justify this assumption.

Note that the condition cannot hold for all time, starting with $F(g_t^1)(C_t - g_t^1) - ag_t^1 > 0$, because if it did then all carbon would end up in the atmosphere. This is a physical impossibility. For an earlier extraction path to lead to a higher temperature path, it is sufficient that the positive feedback condition hold over the positive-extraction phase of the market:

Lemma: *In the dynamic system given by (4), let x^2 represent an earlier extraction path than x^1 , both paths being continuous and terminating at times T^2 and T^1 respectively, with $\int_0^{T^1} x_t^1 dt = \int_0^{T^2} x_t^2 dt$. Let $T = \max(T^1, T^2)$. For common initial conditions (g_0^1, g_0^2) , the path x^2 leads to a higher temperature path provided the positive feedback condition holds on $[0, T]$ for the dynamic systems induced by the extraction paths $x = x^1, x^2$, and all paths in between (according to the "earlier" partial ordering) x^1 and x^2 .*

It follows immediately from the lemma that under the positive feedback condition (6), if a sufficiently small proportion of total fuel deposits have extraction costs above b , then the impact of a successful innovation is a permanently higher temperature path.

Proposition 3 *Under (6):*

(a) *if $[M(b_0) - M(b)]/M(b_0)$ is sufficiently small, where $M(c)$ is the distribution of deposits by extraction cost, then an innovation at any time leads to a higher temperature path.*

(b) *For an arbitrary continuous distribution $M(c)$, there is a time $\hat{t} < T^n$ such that innovation after time \hat{t} leads to a higher temperature path compared to the case of innovation-impossible.*

The availability of innovation always carries with it the chance that no innovation will be discovered early, and the possibility of innovation is worse for global warming than innovation being impossible. The possibility of innovation result of government policy, necessarily adds to the uncertainty of outcomes in a way that exacerbates the worst outcome.

2.3.2 Warming dynamics in the post-extraction phase:

Our formulation of the fossil fuel market, with a backstop technology whether innovation occurs or not, allows a simple characterization of post-extraction dynamics. With no new carbon being ejected into the system in the post-extraction period, $g_t^1 + g_t^2 = g_T^1 + g_T^2 \equiv C$. It follows that for $t > T$, $\dot{g}_t^1 = -\dot{g}_t^2$, from setting $\dot{s}_t = 0$ in (5). Substituting $g_t^2 = C - g_t^1$ into the first equation of (5) allows us to express the post-extraction dynamics of the atmospheric GHG g_t^1 (and therefore the dynamics of global temperature) in a single autonomous differential equation:

$$\dot{g}_t^1 = F(g_t^1)(C - g_t^1) - ag_t^1 \quad (7)$$

for $t \geq T$ with boundary condition given by the exogenous (to this system) value g_T^1 .

The equation (7) offers a particularly clear avenue for examining the effects on the path towards a steady state temperature. The market, under various policies or innovation events, leaves the post-extraction period with different values of C and g_T^1 ; and the impact on the path of the policies can be examined through these two variables.

The function F is of course critical in this examination. The scientific evidence for the feedback effect, reviewed earlier, justifies an assumption that the function is nearly constant at low temperature values, the consensus being that the feedback effects are not problematic at low temperatures; the release of methane gas from the frozen tundra, for example, did not occur at a problematic rate when temperatures were lower. Given the simplicity of our model assumptions, however, an unbounded rate of carbon release would lead to a steady state with no carbon on the earth's surface, a physical impossibility in the real environment. Accordingly, we consider the case where the function F is bounded with the rate a falling between the lower and upper bound.

This leaves us with a sigmoid shape for F as being most natural, as in Figure 3. F leaves \dot{g}_t^1 bounded between two values, v and w . An example of a such a function is the logistic function

$$F(x) = v + \left(\frac{1}{1 + e^{-kx}} \right) (w - v)$$

To understand the resulting dynamics, with $v < a < w$, note that for low values of g^1 , $F(g^1) \approx v$. The differential equation (7) becomes approximately $\dot{g}^1 = v(C - g^1) - ag^1$, which has a unique,

stable steady state at $\dot{g}^1 = 0 \implies g^1 = vC/(v + a)$.¹⁷ Similarly at high values of g^1 , F is approximately w and there is a second steady state at approximately $g^1 = wC/(w + a)$.¹⁸ This second steady state is at a higher level of g^1 , and therefore a higher temperature, since $wC/(w + a) > vC/(v + a)$.

Figure 4 depicts equation (7) with F taking on the logistic functional form, for particular values of the parameter k (which measures the “steepness” of the logistic function at middle values) and values of a and C . The intersections with the coordinate axis are the possible steady states of the system. As always, there is (generically) an odd number of steady states; for the parameter values in this figure, there are three steady states. The middle steady state is unstable. The arrows on the g^1 axis represent the phase space or phase line, which is only one dimension for this simple system.

If extraction costs are homogenous, then there is only one value of C , since all fossil fuels will be exhausted (if $b > c$, the cost of extraction). An earlier extraction is unambiguously harmful and may lead to the higher steady state temperature. A more complete picture of the steady state dynamics in the case of heterogeneous costs is with a phase diagram using the space (g^1, g^2) , keeping in mind that with successful innovation C is reduced. This phase diagram is depicted in Figure 5. Figure 5 has been drawn for a particular choice of a logistic function for F . Note that for $C > C^*$ in the figure, there is only one steady state.¹⁹ An implication that flows immediately from the figure is that a change in the endowment (g_T^1, g_T^2) left for nature by the market (T now being the exhaustion date under any market scenario), may lead to the higher steady state temperature even if the change involves a decrease in C_T . A market leaving the environment at point A in Figure 5, with lower total carbon less than B as a result of successful innovation, will nonetheless lead to a higher steady state temperature because it leaves a higher proportion of the carbon in the atmosphere – as a result, for example, of early innovation combined with strong positive feedback effects in the carbon cycle. An innovation that leads to some substitution from high-extraction-cost carbon fuels to clean energy, and therefore less total carbon injected into the carbon cycle, may nonetheless lead to a higher steady-state temperature because a higher proportion of this endowment, given the positive-feedback during the extraction period, will be

¹⁷For example, if $C = 100$, $v = .01$ and $e = .02$, this yields a steady state value of 33.

¹⁸An approximate value for another locally stable steady state for $w = .06$ and $e = .02$, would be 75.

¹⁹When $F(\cdot)$ is given by

$$F(x) = v + \frac{w - v}{1 + e^{-k(x-50)}}$$

with $v = 0.1$; $w = 0.9$; $e = 0.9$; $k = 0.8$, C^* is approximately 500.

in atmospheric carbon.²⁰

3 Innovation and the Optimal Carbon Tax

In this section, we switch from positive economics to the analysis of optimal policy in the presence of innovation. A growing literature analyzes the dynamic pattern of optimal carbon taxes (See Hoel and Kverndokk 1996). The aim here is more limited. This paper has developed the theme that without a carbon tax, innovation make make global warming worse. We buttress the point in this section by showing that with optimal carbon taxes in place, innovation is always beneficial. An additional result is that the optimal carbon tax in response to innovation is never so high as to offset completely any detrimental impact of innovation. Even in the first best, innovation worsens global warming because the planner, as in any market with externalities, responds to a drop in private costs by tolerating an increase in the externality.

We adopt the objective of maximizing the expected present value of the benefits from energy minus the sum of extraction costs and the costs of global warming. We represent the current cost of global warming as a convex function $D(g^1)$ of the current temperature, setting aside the important lags in this relationship as well as any irreversibilities in damage to the environment. Let $(x_t^n, y_t^n; x_t^d, y_t^d)$ be the consumption of fossil fuel energy and clean energy that a social planner chooses, with superscripts n and d representing the choice of these variables prior to and after discovery of the new clean-energy technology at (random) date τ .

The ex ante optimization problem is:

$$\begin{aligned} \max E_\tau \int_0^\tau e^{-rt} [u(x_t^n + y_t^n) - c(s_t)x_t^n - b_0 y_t^n - D(g_t^1)] dt + \\ e^{-r\tau} \int_\tau^\infty e^{-rt} [u(x_t^d(\tau) + y_t^d(\tau)) - c(s_t)x_t^d(\tau) - b_0 y_t^n(\tau) - D(g_t^1)] dt \end{aligned}$$

subject to

$$\dot{s}_t = x_t \tag{8}$$

$$\dot{g}_t^1 = x_t + F(g_t^1)g_t^2 - ag_t^1 \tag{9}$$

$$\dot{g}_t^2 = -F(g_t^1)g_t^2 + ag_t^1 \tag{10}$$

²⁰The idea that the temperature path will be bumped by innovation from precisely determined steady state to another steady state with a higher temperature is of course an abstraction. The higher temperature steady state represents a “runaway” temperature path due to the positive feedback effects. As a recent *Science* commentary, cited in Weitzman (2009), states, “Once the world has warmed by 4C, conditions will be so different from anything we can observe today ... that it is inherently hard to say where the warming will stop.”

$$x_t, y_t \geq 0 \quad (11)$$

The expectation in the objective is with respect to the exponential distribution on τ . The optimization problem upon discovery of an innovation at date τ (omitting superscripts) is

$$\max \int_{\tau}^{\infty} e^{-rt} [u(x_t + y_t) - c(s_t)x_t - by_t - D(g_t^1)] dt \quad (12)$$

subject to the constraints (8) through (11).

The analysis of steady states in the previous section shows that the social welfare function is discontinuous. At any plan leading to the unstable equilibrium, the social welfare jumps up or down with arbitrarily small changes in the plan. This discontinuity prevents the general application of Debreu's maximum theorem, that a continuous function always reaches a maximum on a closed set, so that formally we are not assured that a social optimum even exists. I simply set this problem aside by assuming that an optimum exists that involves convergence of the environment to a steady state. The optimum is then characterized by first-order conditions.

We can exploit the Markov structure of the problem with a dynamic programming formulation of the optimal policy. Let $V(s_t, g_t^1, g_t^2)$ be the present value of the planner's problem under the optimal solution, as a function of the current state variables, prior to innovation. Let $V^d(s_t, g_t^1, g_t^2)$ be the value function post-innovation. Starting with the post-innovation optimal policy, the Hamilton-Jacobi-Bellman equation is

$$\begin{aligned} rV^d(s, g^1, g^2) &= \max_{x,y} \{u(x+y) - c(s)x - by - D(g^1) \\ &\quad + \frac{\partial V^d}{\partial s} \dot{s} + \frac{\partial V^d}{\partial g^1} \dot{g}^1 + \frac{\partial V^d}{\partial g^2} \dot{g}^2\} \end{aligned}$$

Substituting in the equations of motion (8) through (10) yields

$$\begin{aligned} rV^d(s, g^1, g^2) &= \max_{x,y} \{u(x+y) - c(s)x - by - D(g^1) \\ &\quad + \frac{\partial V^d}{\partial s} x + \frac{\partial V^d}{\partial g^1} [x + F(g^1)g^2 - ag^1] + \frac{\partial V^d}{\partial g^2} [-F(g^1)g^2 + ag^1]\} \end{aligned} \quad (13)$$

The first-order conditions for the ex post problem are

$$u'(x+y) - c(s) + \left[\frac{\partial V^d}{\partial s} + \frac{\partial V^d}{\partial g^1} \right] \leq 0 \quad (= 0 \text{ when } x > 0) \quad (14)$$

$$u'(x+y) - b \leq 0 \quad (= 0 \text{ when } y > 0) \quad (15)$$

The planner can implement this optimal policy with a competitive market and a tax per unit of carbon given by $\theta = -\partial V^d / \partial g^1$. To show this, note first the competitive price

b for clean energy clears the market at the optimal quantity when $y > 0$, from (15). Second, the buyer's price for the fossil fuel will satisfy $u'(x + y) = p$ when $x > 0$ and on the sellers' side of the market we know that $p = c(s) + \pi + \theta$ where π is the scarcity rent. The market will therefore clear at the optimal values for x and y , under the prescribed tax rate, provided that the scarcity rent, $\pi = p - c(s) - \theta$, is equal to $\partial V^d / \partial s$.

Letting x^* and y^* refer to the optimal choices of x and y , (13) yields (omitting arguments of V^d)

$$rV^d = u(x^*, y^*) - c(s)x^* - by^* - D(g_1) + V_s^d x^* + V_{g^1}^d [x^* + F(g^1)g^2 - ag^1] + V_{g^2}^d [-F(g^1)g^2 + ag^1]$$

Differentiating this with respect to s yields

$$rV_s^d = -c'(s)x^* + V_{ss}^d x^* + V_{g^1 s}^d g^1 + V_{g^2 s}^d g^2 \quad (16)$$

Note that $dV_s^d / dt = V_{ss}^d \dot{x}^* + V_{g^1 s}^d \dot{g}^1 + V_{g^2 s}^d \dot{g}^2$. Therefore (16) implies

$$\frac{dV_s^d}{dt} - rV_s^d = c'(s)x^* \quad (17)$$

Equation (17) can be obtained directly as the co-state equation in a maximum-principle approach to the optimization problem. As a differential equation in V_s^d , (17) is equivalent to the differential equation (1) characterizing the competitive market value of $\pi = p - c(s) - \theta$ (when (1) is extended to include the carbon tax). Furthermore, the boundary conditions to the social optimum and the market equilibrium are equivalent: $\lim_{s \rightarrow c^{-1}(b)} V_s^d = 0$ and $\lim_{s \rightarrow c^{-1}(b)} \pi = 0$, proving that if a competitive price path is set to ensure that $\pi = V_s^d$, the competitive equilibrium conditions are satisfied.²¹ That is, the competitive market implements the social optimum with the carbon tax $\theta = -\partial V^d / \partial g^1$, post-innovation.

The HJB equation for the optimality problem ex ante (prior to discovery) is the following. (For clarity, this equation is derived from first principles in Appendix C):

$$\begin{aligned} rV(s, g^1, g^2) = & \max_{x,y} \{ [u(x+y) - c(s)x - b_0y - D(g^1)] + \rho(V^d(s, g^1, g^2) - V(s, g^1, g^2)) \\ & + [\frac{\partial V}{\partial s} \dot{s} + \frac{\partial V}{\partial g^1} \dot{g}^1 + \frac{\partial V}{\partial g^2} \dot{g}^2] \} \end{aligned} \quad (18)$$

From the equations of motion for s , g^1 and g^2 , (18) can be rewritten

$$rV(s, g^1, g^2) = \max_{x,y} \{ [u(x+y) - c(s)x - b_0y - D(g^1)] + \rho(V^d(s, g^1, g^2) - V(s, g^1, g^2)) \} \quad (19)$$

²¹The boundary conditions are limits as s approaches the total amount that will be exhausted as $t \rightarrow \infty$ because in the optimum, unlike the competitive market without taxes, x^* does not in general reach zero. Hoel and Kverndokk (1996) demonstrate this in a somewhat simpler model.

$$+[\frac{\partial V}{\partial s}x + \frac{\partial V}{\partial g^1}[x + F(g^1)g^2 - ag^1] + \frac{\partial V}{\partial g}[-F(g^1)g^2 + ag^1]]\}$$

The first order conditions for the ex ante problem are

$$u'(x + y) - c(s) + \frac{\partial V}{\partial s} + \frac{\partial V}{\partial g^1} = 0 \quad (20)$$

$$u'(x + y) - b_0 = 0 \quad \text{when } y > 0 \quad (21)$$

The proof provided above that the equilibrium post-innovation can be implemented with a carbon tax θ has a parallel in the ex ante case as well, showing that overall a carbon tax implements the socially optimal plan $\{x(s, g^1, g^2), y(s, g^1, g^2)\}$.

The optimal carbon tax $\theta = -V_{g^1}^n$ may decrease with the event of innovation. For example, if innovation occurs at a date t where s_t is very close to $c^{-1}(b)$ then the shadow cost of emissions drops because future carbon emissions are then known to be much lower than the expected emissions just prior to innovation. On the other hand, if innovation is very early and the positive-feedback effect in the carbon cycle is strong, then the shadow price and the optimal tax will increase.

The optimal carbon tax will never rise to offset completely the impact on consumer price of the drop in the opportunity cost of energy, however. Evaluated at the ex ante optimal plan, (x^{n*}, y^{n*}) , which solves (20), the first-order condition for the ex post plan, (14) must be positive: the final term $\partial V^d / \partial g^1$ equals $\partial V / \partial g^1$ from (20) since the future path of damage from global warming is unchanged. The first two terms, $u'(x + y) - c(s)$, are also identical between the two first-order conditions evaluated at (x^{n*}, y^{n*}) , but $\partial V^d / \partial s < \partial V / \partial s$ with the innovation making the left hand side of (14) smaller than the left hand side of (20). A dominating strategy to leaving the plan unchanged at (x^{n*}, y^{n*}) is therefore a small increase in x . The small increase in x can only be achieved with a decrease in p . Finally, with carbon taxes available as an instrument, innovation necessarily increases social welfare: the (suboptimal) plan of leaving (x, y) unchanged by imposing high carbon taxes leads to identical damages, identical consumer benefits and with the lower cost of extraction, higher social welfare. The following proposition summarizes the results of this section

Proposition 4 *When carbon taxes are available as an instrument, a successful innovation always raises social welfare. The carbon taxes allow the implementation of the social optimum (x^*, y^*) , which involves an increase in current emissions in the event of innovation.*

4 Conclusion

This paper has explored a consequence of clean-energy innovation that has received too little attention in climate change policy literature. Without carbon pricing, innovation can easily make the global problem worse, through its impact on endogenous prices of fossil fuels. The perverse warming effect of innovation is the greater intensity of fossil fuel consumption and release of carbon emissions at an earlier time. The beneficial effect is the future displacement of high-extraction-cost fossil fuels with newly discovered clean energy. When feedback effects in the carbon-temperature dynamics are strong and the wait for beneficial displacement effects is long, the perverse effect will dominate.

Evidence is consistent with this argument being important. Hotelling rents are high in crude oil markets, with a price in the range of 100 dollars per barrel, and extraction costs of conventional crude less than 30 dollars. The International Energy Agency estimated recently that more than 4 trillion barrels of oil are available at extraction costs of 60 dollars per barrel or less (about 1 trillion have been extracted to date).²² In other words, the stock of oil that would be consumed more rapidly and earlier in the event of a significant innovation in clean energy will last many decades. The displacement of high-cost crude by clean energy is in the distant future.

Two of the main instruments for global warming policy are carbon pricing, through taxes or a cap-and-trade system, and government support of innovation. As a naked instrument, support of innovation can make things worse. Carbon pricing that is *responsive* to innovation, however, will offset the dark side of clean-energy innovation. Where the net impact of innovation is detrimental, a responsive tax will magnify the cost advantage of any newly developed clean-energy technology by raising carbon prices. The political challenges with the implementation of a policy that magnifies the negative impact on oil and gas industry profits of success in competing technologies are obvious. But the message of this paper is that while the route of subsidies may be politically easier than carbon pricing, it is not a solution at all. Innovation may make global warming worse.²³

²²IEA (2005).; see also Stern (2006) at 212. These are 2004 U.S. dollars. Current world consumption of oil is about 30 billion barrels per year.

²³ An alternative to support of innovation or cost subsidies to clean energy technology is the support of alternative energy with minimum share targets, e.g. 20 percent of energy to be wind or solar within 30 years. This type of policy has two effects. As the policy is implemented, the decrease in demand for fossil fuels helps global warming. But the anticipation of the (announced) implementation of the policy in the future has a detrimental impact on emissions today, as Strand (2007) and Sinn (2008) have emphasized. Given that the

Important aspects of the interaction of innovation and climate change dynamics have been set aside in this paper. The carbon cycle is more complex than the risibly simple dynamics assumed here. In terms of economics, the full range of relevant technologies and innovation include not just clean energy, but technologies to mitigate carbon emissions (such as carbon sequestration or cleaner automobile engines); demand-side innovation to make more efficient use of energy; and even technology to directly mitigate global warming (such as injecting reflective particles in the atmosphere).

More general changes in technology over time can add to the perverse effect clean energy innovation. An exogenous rate of improvement in mitigation technology such as carbon sequestration, for example, adds to the opportunity cost of the earlier extraction of fossil fuels induced by innovation because of higher rates of carbon emission prior to the development of mitigation technology. This paper is simply one step towards integrating innovation into the economics of global warming policy.

market share of wind and solar are so small (less than 3% of world energy supply currently), the detrimental effect dominates in terms of carbon emissions in the short run.

Appendices

Appendix A: Derivation of Equation (2)

I derive this equation as follows. For a price-taking firm to be indifferent between extracting at t and at $t + dt$, the present value of profits must be equal:

$$\begin{aligned}
 (1 + rdt) [p_t^n - c(s_t^n)] &= \rho dt \left[p_{t+dt}^d - c \left(s_t^d(t) \right) \right] + (1 - \rho dt) [p_{t+dt}^n - c(s_t^n)] \\
 &= \rho dt [p_{t+dt}^d(t) - c(s_{t+dt}^d(t))] + (1 - \rho dt) [p_{t+dt}^n - c(s_{t+dt}^n)] \\
 &\quad + \rho dt \left[\frac{\partial c(s_t^d(t))}{\partial t} dt \right] + (1 - \rho dt) \left[\frac{\partial c(s_t^n)}{\partial t} dt \right]
 \end{aligned}$$

To a first-order approximation:

$$\begin{aligned}
 (1 + rdt) [p_t^n - c(s_t^n)] &= \rho dt [p_t^d(t) - c(s_t^d(t))] + \frac{\partial (p_t^d(t) - c(s_t^d(t)))}{\partial t} dt + (1 - \rho dt) [p_t^n - c(s_t^n)] + \frac{\partial (p_t^n - c(s_t^n))}{\partial t} dt \\
 &\quad + \rho dt \left[\frac{\partial c(s_t^d(t))}{\partial t} dt \right] + (1 - \rho dt) \left[\frac{\partial c(s_t^n)}{\partial t} dt \right]
 \end{aligned}$$

Subtracting $[p_t^n - c(s_t^n)]$ from both sides, leads to

$$\begin{aligned}
 (rdt) [p_t^n - c(s_t^n)] &= \rho dt [p_t^d(t) - c(s_t^d(t))] + \frac{\partial (p_t^d(t) - c(s_t^d(t)))}{\partial t} dt - \rho dt [p_t^n - c(s_t^n)] + (1 - \rho dt) \frac{\partial (p_t^n - c(s_t^n))}{\partial t} dt \\
 &\quad + \rho dt \left[\frac{\partial c(s_t^d(t))}{\partial t} dt \right] + (1 - \rho dt) \left[\frac{\partial c(s_t^n)}{\partial t} dt \right]
 \end{aligned}$$

Dividing both sides by dt and taking the limit as $dt \rightarrow 0$ yields

$$r[p_t^n - c(s_t^n)] = \rho [p_t^d(t) - c(s_t^d(t)) - (p_t^n - c(s_t^n))] + \frac{\partial (p_t^n - c(s_t^n))}{\partial t} + c'(s_t^n) \dot{s}_t^n(t)$$

which is (2). This can be re-expressed as

$$\frac{\frac{\partial}{\partial t} (p_t^n - c(s_t^n))}{p_t^n - c(s_t^n)} = r + \rho \left(1 - \frac{p_t^d(t) - c(s_t^d(t))}{p_t^n - c(s_t^n)} \right) - \frac{c'(s_t^n) \dot{s}_t^n(t)}{p_t^n - c(s_t^n)}$$

or (since $c(s_t^d(t)) = c(s_t^n)$)

$$\frac{\frac{\partial}{\partial t} (p_t^n - c(s_t^n))}{p_t^n - c(s_t^n)} = r - \frac{c'(s_t^n) \dot{s}_t^n(t)}{p_t^n - c(s_t^n)} + \rho \left(\frac{p_t^n - p_t^d(t)}{p_t^n - c(s_t^n)} \right)$$

Appendix B:

Lemma: *In the dynamic system given by (4), let \tilde{x} represent an earlier extraction path than x , both paths being continuous and terminating at times \tilde{T} and T respectively, with $\int_0^{\tilde{T}} \tilde{x} dt = \int_0^T x dt$. Let $\hat{T} = \max(\tilde{T}, T)$. For common initial conditions (g_0^1, g_0^2) , the path \tilde{x} leads to a higher temperature path provided the positive feedback condition holds on $[0, \hat{T}]$ for the dynamic systems induced by the extraction paths $x = \tilde{x}$ and x , and all paths in between (according to the “earlier” partial ordering) \tilde{x} and x .*

Proof. Consider a function $\Delta = 1$ on $(t_1, t_1 + \theta)$ and -1 on $(t_2, t_2 + \theta)$, for $t_1 < t_2 < T$ and θ (vanishingly small). Since the paths x and \tilde{x} are continuous and bounded, the paths are integrable. It therefore suffices to show that a move from x to $x + \Delta$ results in a higher temperature path under (6) since the difference between x and \tilde{x} can be constructed from an infinite weighted sum of such differences. (We must also recognize that only weak inequalities are preserved in the limit of the infinite sum.) Let \tilde{g}^1 represent the path of atmospheric carbon under the path $x + \Delta$ and g^1 the path of atmospheric carbon under the path x . The right hand derivatives of g^1 and \tilde{g}^1 evaluated at t_1 satisfy $\partial \tilde{g}^1 / \partial t^+ = \partial g^1 / \partial t + 1 > \partial g^1 / \partial t$. It follows that there exists a value $\hat{\theta}$ such that for every $t \in (t_1, t_1 + \theta)$, $\tilde{g}_t^1 > g_t^1$. We also have that for each such t , $\tilde{C}_t > C_t$ since the total amount of carbon injected into the biosphere is greater under $x + \Delta$ than under x . From this and (6) we have for $t \in (t_1, t_1 + \theta)$, an inequality in the net flows of carbon to the atmosphere under x and $x + \Delta$:

$$F(\tilde{g}_t^1)\tilde{g}_t^2 - a\tilde{g}_t^1 = F(\tilde{g}_t^1)(\tilde{C}_t - \tilde{g}_t^1) - a\tilde{g}_t^1 > F(\tilde{g}_t^1)(C_t - \tilde{g}_t^1) - a\tilde{g}_t^1 > F(g_t^1)(C_t - g_t^1) - ag_t^1 = F(g_t^1)g_t^2 - ag_t^1 \quad (22)$$

Consider the set $\Upsilon = \{t \leq \hat{T} \mid \tilde{g}_t^1 < g_t^1 \text{ or } d\tilde{g}_t^1/dt = F(\tilde{g}_t^1)\tilde{g}_t^2 - a\tilde{g}_t^1 < F(g_t^1)g_t^2 - ag_t^1 = dg_t^1/dt\}$, i.e. the set for which at least one of the two inequalities we have established for $t \in (t_1, t_1 + \theta)$ is violated. We show that the set Υ is empty, from which it follows that $\tilde{g}_t^1 > g_t^1$ for all t . Suppose that Υ is not empty. From the continuity of the functions defining Υ , $t^* \equiv \min \Upsilon$ exists and at least one of the following two equalities holds: (a) $\tilde{g}_{t^*}^1 = g_{t^*}^1$; (b) $F(\tilde{g}_{t^*}^1)(\tilde{C} - \tilde{g}_{t^*}^1) - a\tilde{g}_{t^*}^1 = F(g_{t^*}^1)(C - g_{t^*}^1) - ag_{t^*}^1$. But (a) cannot hold since for all $t < t^*$, $\tilde{g}_t^1 > g_t^1$ and $d\tilde{g}_t^1/dt > dg_t^1/dt$. Suppose that (b) holds. The function $F(g)(C-g) - ag$ is increasing both in C and, by hypothesis, in g ; and furthermore, $\tilde{C}_t \geq C_t$ for all t by construction since the carbon emitted is weakly higher under $x + \Delta$. It follows that the first term on the left hand side of (b) is no less than the first term on the right hand side. From the supposition of (b) we then have that $\tilde{g}_{t^*}^1 \leq g_{t^*}^1$. The strict inequality $\tilde{g}_{t^*}^1 < g_{t^*}^1$ cannot hold by definition of t^* and continuity; the equality $\tilde{g}_{t^*}^1 = g_{t^*}^1$ is (a) and has already been ruled out.

It follows that $\tilde{g}_t^1 > g_t^1$ for all t in (t_1, \hat{T}) . It remains to show that $\tilde{g}_t^1 \geq g_t^1$ for $t > T$. For $t > T$, the contribution of fossil fuels to g_t^1 is zero, i.e. $\dot{s} = 0$ and the system (5) reduces to a single differential equation $\dot{g}_t^1 = F(g_t^1)(C - g_t^1) - ag_t^1$, with initial value g_T^1 , where C is the constant value of $g_t^1 + g_t^2 = g_T^1 + g_T^2$. It is easily verified from this equation that the inequality in the initial values, $\tilde{g}_T^1 > g_T^1$, the fact that $\dot{g}_T^1 > 0$ and $\tilde{C} = C$ imply that $\tilde{g}_t^1 \geq g_t^1$ for $t > T$: suppose to the contrary that there is a crossing point, t^* , at which $\tilde{g}_{t^*}^1 = g_{t^*}^1$. Starting at t^* and solving the autonomous ODE backwards yields $\tilde{g}_T^1 = g_T^1$, contradicting $\tilde{g}_T^1 > g_T^1$.

Therefore $\tilde{g}_t^1 > g_t^1$ for all t . Constructing \tilde{x} as the limit of an infinite weighted sum of such differences preserves this as a weak inequality. The partial order of a "higher" temperature path over (t_1, \hat{T}) simply requires that we rule out equality of the temperature paths, which is trivial. ■

Appendix C: Derivation of the HJB equation for the ex ante (prior to discovery) social optimum, equation (18):

Bellman's principle of optimality implies that the following first-order approximation, for small dt :

$$\begin{aligned} V(s_t, g_t^1, g_t^2) &= \max_{x,y} \{ [u(x_t + y_t) - c(s_t)x_t - b_0y_t - D(g_t^1)]dt \\ &\quad + \frac{1}{1 + rdt} [(1 - \rho dt)V(s_{t+dt}, g_{t+dt}^1, g_{t+dt}^2) + \rho dt V^d(s_{t+dt}, g_{t+dt}^1, g_{t+dt}^2)] \} \end{aligned}$$

Additional first order approximations yield

$$V^d(s_{t+dt}, g_{t+dt}^1, g_{t+dt}^2) = V^d(s_t, g_t^1, g_t^2) + \frac{\partial V^d}{\partial s} \dot{s} dt + \frac{\partial V^d}{\partial g^1} \dot{g}^1 dt + \frac{\partial V^d}{\partial g^2} \dot{g}^2 dt$$

$$V(s_{t+dt}, g_{t+dt}^1, g_{t+dt}^2) = V(s_t, g_t^1, g_t^2) + \frac{\partial V}{\partial s} \dot{s} dt + \frac{\partial V}{\partial g^1} \dot{g}^1 dt + \frac{\partial V}{\partial g^2} \dot{g}^2 dt$$

Therefore

$$\begin{aligned} (1 + rdt)V(s_t, g_t^1, g_t^2) &= \max_{x,y} \{ (1 + rdt)[u(x_t + y_t) - c(s_t)x_t - b_0y_t - D(g_t^1)]dt \\ &\quad + \rho dt [V^d(s_t, g_t^1, g_t^2) + \frac{\partial V^d}{\partial s} \dot{s} dt + \frac{\partial V^d}{\partial g^1} \dot{g}^1 dt + \frac{\partial V^d}{\partial g^2} \dot{g}^2 dt] \\ &\quad + (1 - \rho dt)[V(s_t, g_t^1, g_t^2) + \frac{\partial V}{\partial s} \dot{s} dt + \frac{\partial V}{\partial g^1} \dot{g}^1 dt + \frac{\partial V}{\partial g^2} \dot{g}^2 dt] \} \end{aligned}$$

Subtracting $V(s_t, g_t^1, g_t^2)$ from both sides

$$\begin{aligned}
 rdtV(s_t, g_t^1, g_t^2) &= \max_{x,y} \{ (1 + rdt)[u(x_t + y_t) - c(s_t)x_t - b_0y_t - D(g_t^1)]dt \\
 &\quad + \rho dt[V^d(s_t, g_t^1, g_t^2) + \frac{\partial V^d}{\partial s} s^d dt + \frac{\partial V^d}{\partial g^1} g^1 dt + \frac{\partial V^d}{\partial g} g^2 dt] \\
 &\quad - \rho dtV(s_t, g_t^1, g_t^2) + (1 - \rho dt)[\frac{\partial V}{\partial s} s dt + \frac{\partial V}{\partial g^1} g^1 dt + \frac{\partial V}{\partial g} g^2 dt]
 \end{aligned}$$

Dividing both sides by dt , and taking limit, we get equation (18).

References

- Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, David Hemous (2011), "The Environment and Directed Technical Change," NBER working paper, forthcoming, *American Economic Review*
- Allen, Myles R., and David J. Frame, "Call Off the Quest," *Science* 318 (October 26, 2007), 582–583.
- Antweiler, Werner (2010), "Mitigation plus Adaption: What is the Optimal Climate Change Policy Mix?" Sauder School of Business Working Paper.
- Peter M. Cox, Richard A. Betts, Chris D. Jones, Steven A. Spall & Ian J. Totterdell (2000) "Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model," *Nature*, November 2000.
- Dasgupta, P., Stiglitz, J., "Resource Depletion under Technological Uncertainty," *Econometrica* 49, 1 (Jan 1981): 85-104.
- Devarajan, S. and A.C. Fisher (1981), "Hotelling's 'Economics of Exhaustible Resources': Fifty Years Later," *Journal of Economic Literature*, Vol. 19(1): 65-73.
- Gallini, N., Lewis, T., Ware, R., "Strategic Timing and Pricing of a Substitute in a Cartelized Resource Market," *Canadian Journal of Economics*, 1983.
- Hayward, Steven, Munro, Mark, Nordhaus, Ted and Shellenberger, Michael (2010) *Post-Partisan Power* (American Enterprise Institute; Brookings Institution; Breakthrough Institute).

- Heal, Geoffrey, "The Relationship between Price and Extraction Cost for a Resource with a Backstop Technology," *Rand Journal of Economics*, Autumn 1976: 371-378.
- Hoel, Micheal (2008), "Bush meets Hotelling: The Effects of Improved Renewable Energy Technology on Greenhouse Gas Emissions," University of Oslo W.P.
- Hoel, Michael and Snorre Kvemdokk (1996), "Depletion of fossil fuels and the impacts of global warming", *Resource and Energy Economics*: 115-136
- Hoel, Michael (2012), "Carbon Taxes and the Green Paradox," forthcoming, R. Hahn and U. Ulph (eds.): "Climate Change and Common Sense: Essays in Honour of Tom Schelling", Oxford University Press.
- Hotelling, Harold (1931), "The Economics of Exhaustible Resources," *Journal of Political Economy*, vol. 39: 137-175.
- Hsu, Shi-Ling (2009), "Nine Reasons to Adopt a Carbon Tax," SSRN Working Paper.
- Intergovernmental Panel on Climate Change. "Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC."
- International Energy Association (2010), "Resources to Reserves".
- Nordhaus, William (2007), "The Challenge of Global Warming: Economic Models and Environmental Policy," April 4, 2007, Yale University
- Nordhaus, William (2008), *A Question of Balance*, Yale University Press.
- Nordhaus, William (2011), "Estimates of the Social Cost of Carbon: Background and Results from the RICE-2011 Model," Cowles Foundation Working Paper (October)
- Raupach, M.R. et al. (2007). "Global and regional drivers of accelerating CO2 emissions". *Proc. Natl. Acad. Sci. U.S.A.* 104 (24): 10288–93
- Reinganum, Jennifer F. (1989), "Chapter 14 The timing of innovation: Research, development, and diffusion", In: Richard Schmalensee and Robert Willig, Editor(s), *Handbook of Industrial Organization*, Elsevier, Volume 1:849-908.
- Revkin, Andrew, "Bill Gates on R&D, a Carbon Tax and China's Climate Role," Reuters, Sept.2, 2010

- Sinclair, P.J.N. (1992), "High Does Nothing and Rising is Worse: Carbon Taxes Should Keep Falling to Cut Harmful Emissions', Manchester School, 60, 41-52.
- Sinclair, P.J.N. (1994), "On the Optimum Trend of Fossil Fuel Taxation:" Oxford Economic Papers, Vol. 46, (Oct.1994), pp. 869-877
- Sinn, H.W. (2008), "Public Policies against Global Warming: A Supply Side Approach," *International Tax and Public Finance* 15, 2008, p. 360–394.
- Stern, Nicholas (2006), *The Economics of Climate Change*, Cambridge University Press, 2006
- Strand, J. (2007): Technology Treaties and Fossil Fuels Extraction, *The Energy Journal* 28(4), 129-142.
- Ulph, A. and Ulph, D. (1994). 'The Optimal Time Path of a Carbon Tax', Oxford Economic Papers, 46, 857-868.
- Usawa (2003), *Economic Theory and Global Warming*, Cambridge University Press.
- Weitzman, Martin L.,(2009) "On Modelling and Interpreting the Economics of Catastrophic Climate Change," *Review of Economics and Statistics*, 19 (1) (February 2009):1-19.

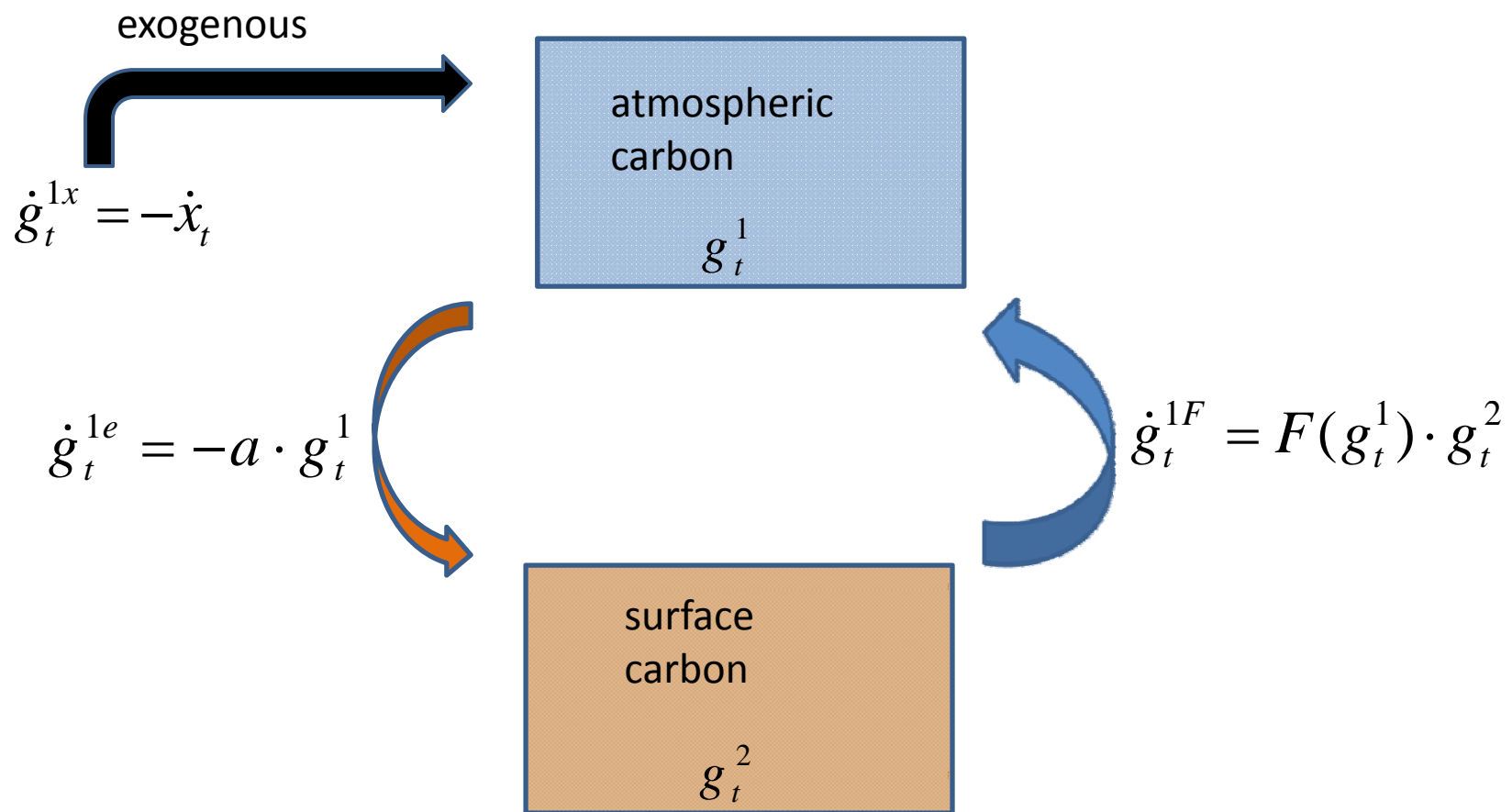


Figure 1: Assumptions on the Carbon Cycle

$$\dot{g}_t^1 = \dot{g}_t^{1x} + \dot{g}_t^{1e} + \dot{g}_t^{1F}$$

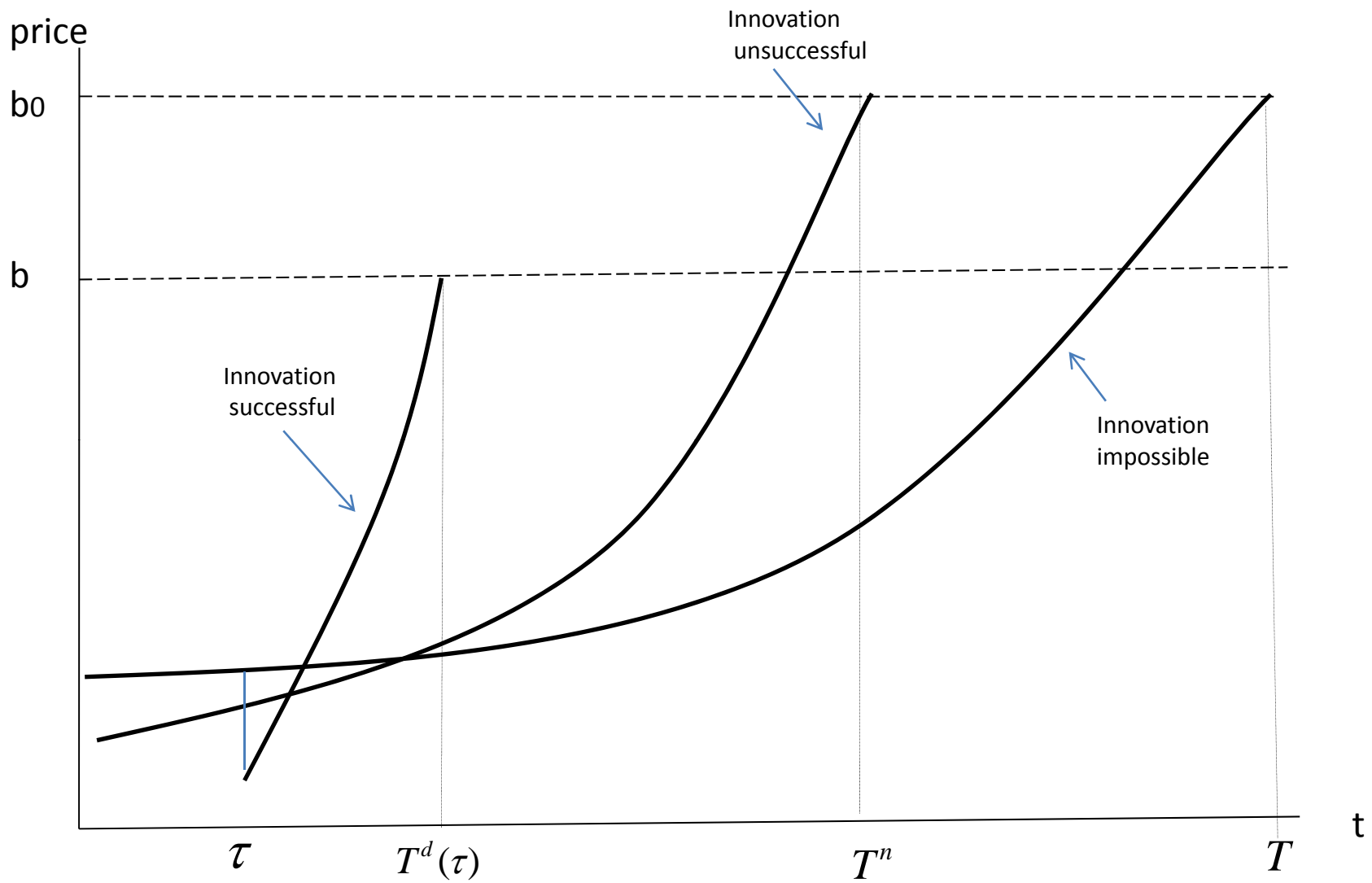


Figure 2: Fossil Fuel Price Paths for Three Cases:

- innovation impossible
- innovation unsuccessful;
- innovation successful at date τ

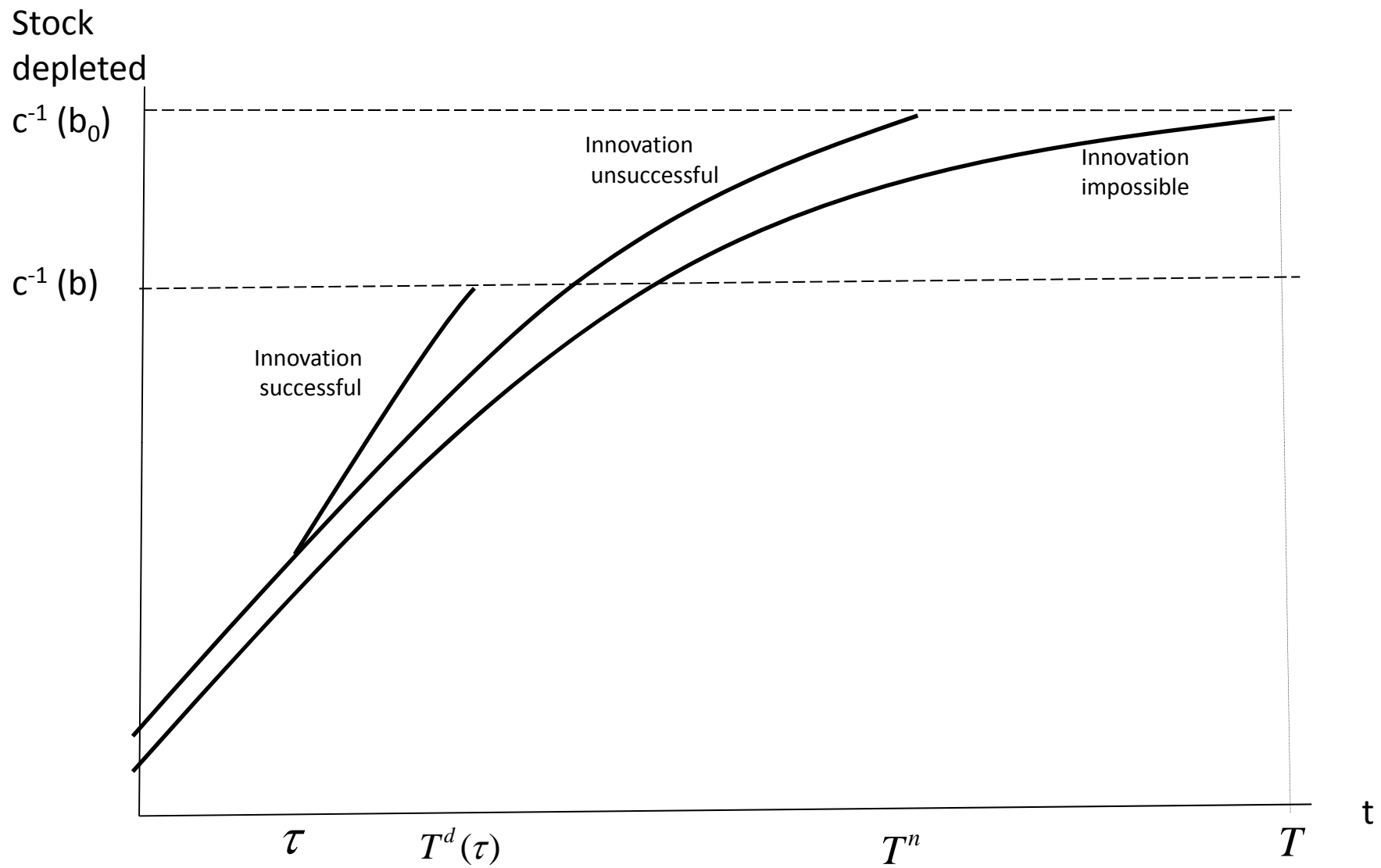


Figure 3: Fossil Fuel Cumulative Exhaustion Paths

- innovation impossible
- innovation unsuccessful
- innovation successful at date τ

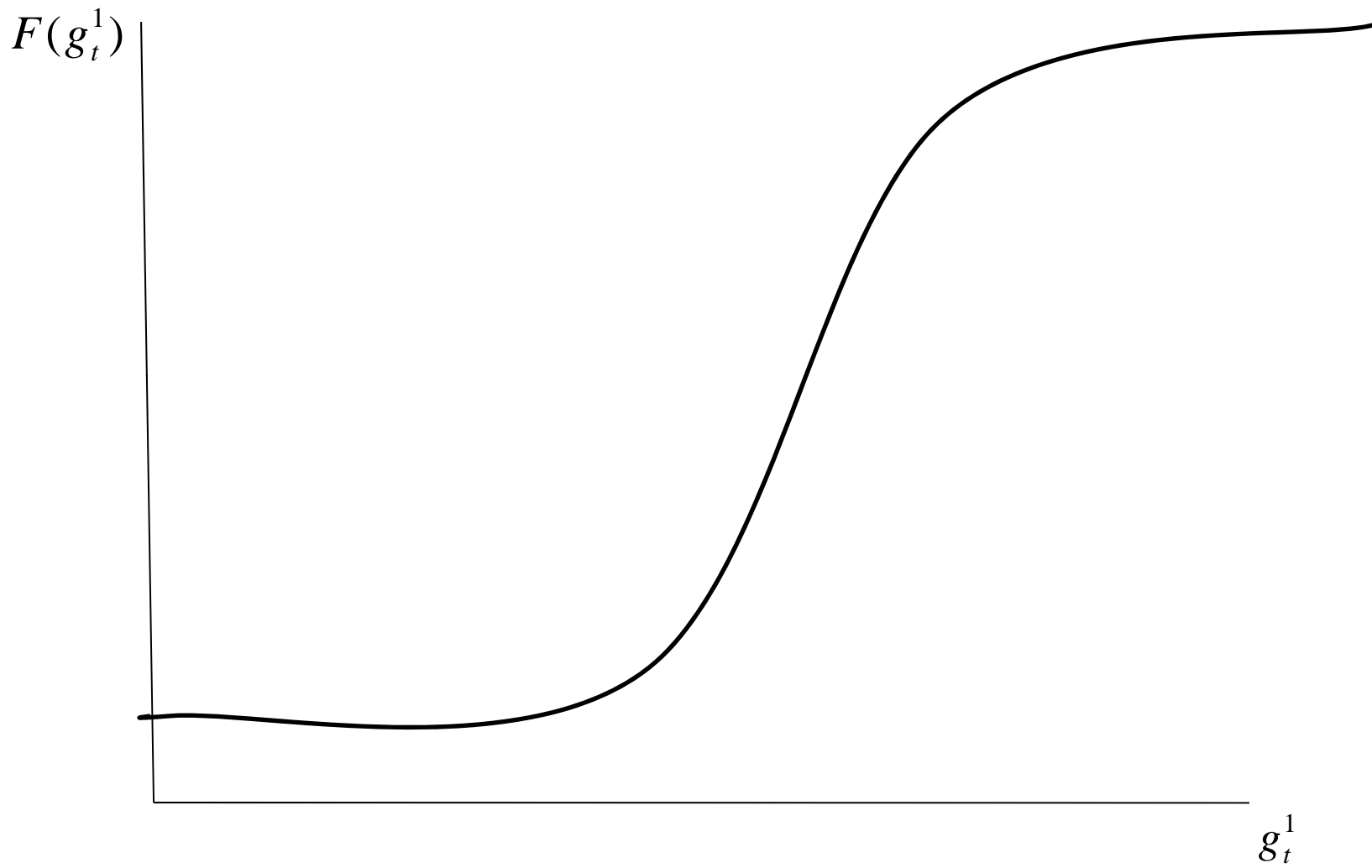


Figure 4: Sigmoid shape for $F(g_t^1)$

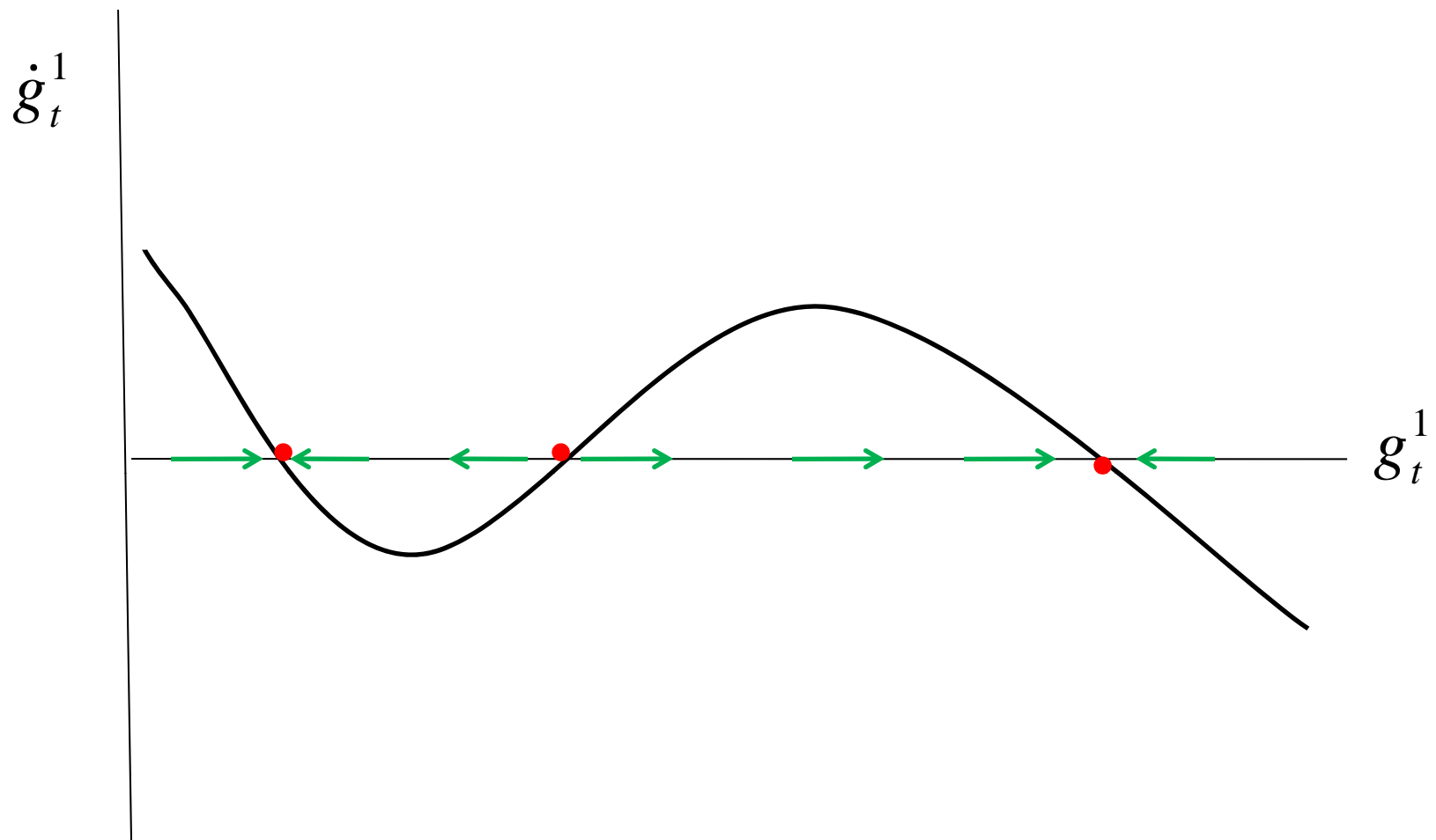


Figure 5: Steady states in g_t^1 for fixed C.

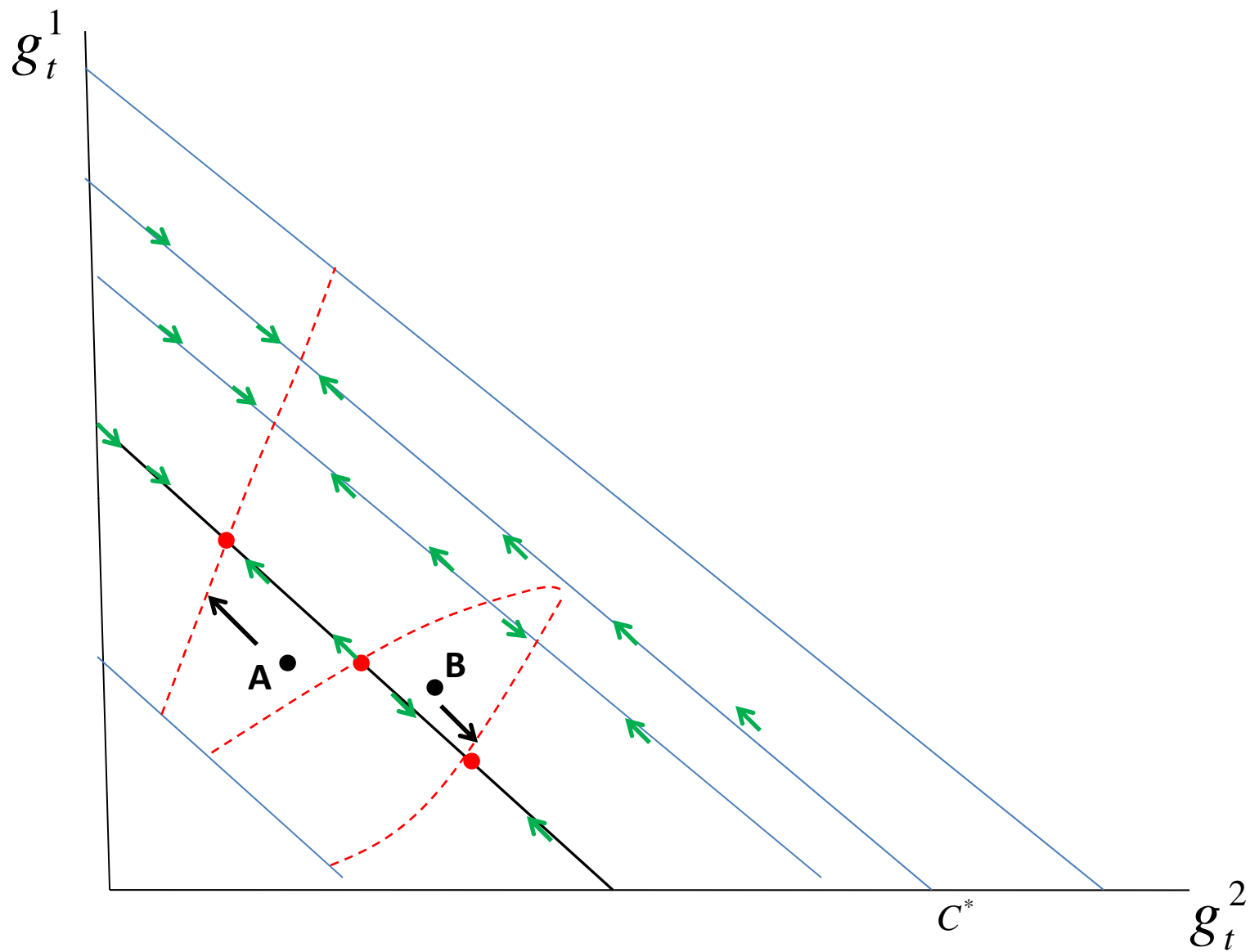


Figure 6: Steady states, for various levels of $C_T = g_T^1 + g_T^2$