

Regulation of markets with sluggish supply structures

Halvor Briseid Storrøsten



Working Paper 02/2018



CREE – Oslo Centre for Research on Environmentally Friendly Energy
acknowledges financial support from
The Research Council of Norway, University of Oslo and user partners.

ISBN: 978-82-7988-252-7

ISSN: 1892-9680

<http://www.cree.uio.no>

Abstract in Norwegian:

- Working Paper 02/2018

Det grønne skiftet: regulering og omstillingsdynamikk

Halvor Briseid Storrøsten

Bedrifter bruker tid til å tilpasse seg nye reguleringer som krever arbeidere med ny kompetanse eller utskiftning av maskiner og bygninger. Dette essayet presenterer en dynamisk modell for omstilling av elektrisitetssektoren på veien mot lavutslippssamfunnet. Modellen viser at annonsering av nye utslippsskatter har tre effekter på utslipp allerede før skattene innføres:

1. Redusert etterspørsel etter fossilt brensel. De annonserte skattene øker kostnaden ved å forbrenne kull og gass i fremtiden. Dette gjør det mindre gunstig å vedlikeholde eller investere i fossile varmekraftverk og senker dermed etterspørselen etter fossile brenslers.
2. Økt tilbud av elektrisitet fra relativt rene energikilder. Utslippsskattene vil gjøre fossil kraftproduksjon dyrere i fremtiden, hvilket innebærer at relativt ren energi blir mer konkurransedyktig. Dette gjør investeringer i blant annet fornybar energi mer attraktivt. Det økte tilbudet av ren elektrisitet reduserer konsumet av elektrisitet fra utslippsintensive varmekraftverk.
3. Økt produksjon av begrensede fossile ressurser som olje og gass. Dette skjer fordi de fremtidige skattene reduserer verdien i å spare ressursene for senere produksjon. Ergo er det mer gunstig å utvinne mer nå. Denne mekanismen refereres gjerne til som det «grønne paradokset».

Mens (1) og (2) er etterspørselssideeffekter som bidrar til å redusere utslippene, er (3) en tilbudssideeffekt som drar i retning av økte utslipp. Fra et teoretisk ståsted er det dermed tvetydig om annonsering av fremtidige skatter vil øke eller senke dagens utslipp. Numeriske simuleringer gir imidlertid sterke indikasjoner på at (1) og (2) dominerer (3), dvs. at utslippene vil falle.

Når det gjelder optimal politikk viser analysen at fremtidige utslippsreduksjoner bør møtes med handling allerede i dag. Dette fordi det tar tid å bygge opp produksjonskapasitet i ren energi, og ned produksjonskapasiteten i eksisterende forurensende kraftverk. Videre er det ikke optimalt å holde utslippsskatten lav i starten for å gi bedrifter tid til å tilpasse seg fremtidig regulering.

Analysen viser at virkemiddelbruken på veien mot lavutslippssamfunnet bør annonseres tydelig og være forutsigbar, da annonseringen i seg selv har effekt og stor verdi. Et effektivt grønt skifte vil kreve en kombinasjon av skatter/subsidier på investeringer i tillegg til skattlegging av utslipp i en overgangsfase dersom bedrifter diskonterer fremtiden hardere enn det som er optimalt fra et velferdsperspektiv.

Regulation of markets with sluggish supply structures

Halvor Briseid Storrøsten*

February 13, 2018

Abstract

This paper examines regulation in the presence of convex investment costs. I find strong analytical and numerical evidence that announcement of future emission taxes will reduce current emissions; i.e., emissions decline before the tax is implemented. The results do not support the hypothesis that taxes should be gradually phased in to avoid excess adjustment costs. On the contrary, lower emissions in the future may require decisive action today, and the lion share of abatement effort occurs early on along the socially optimal time trajectory. The theory is complemented with a numerical model of the U.S. electricity market.

Keywords: regulation, capacity constraints, investment, exhaustible resources, climate change.

JEL classification: H21, H23, Q35, Q41, Q54.

1 Introduction

A power plant or vehicle may operate for decades before it is obsolete. Consequently, adaptation to new regulatory policies will be sluggish. In this paper, I examine regulation and transition dynamics in the presence of adjustment costs caused by convex investment costs.

I first examine the dynamic effects of increasing (or introducing) future emission taxes. Then I derive the socially optimal time trajectory and the tax that can implement it in competitive equilibrium. The results are relevant for regulation of any activity that involves capacity constraints and

*Statistics Norway, Research Department. E-mail: hbs@ssb.no

investment decisions. To fix thoughts, I use regulation of the electric power industry sector as an example throughout this paper. In this setting, convex investment costs entail that the cost of reducing greenhouse gas (GHG) emissions increases with the speed of emission reductions.

Anticipated future emission taxes are shown to have two key dynamic effects on emissions from electricity generation. First, future taxes increase the future cost of combusting fossil fuels. This reduces the profitability of upkeep and investment in emission intensive power plants, which again reduces the demand for fossil fuels. Second, future emission taxes increase future residual demand for low emission electricity. This increases the profitability of investment in low emission power plants, which again increases the electricity supply from these plants to the market. This reduces the equilibrium consumption of fossil fuels. Because of convex investment costs, it is cost efficient for the firms to begin adaptation to anticipated future emission taxes at once. Therefore, these two mechanisms start to operate as soon as the tax announcement has occurred.

Regarding optimal taxation, the regulator faces a trade-off in the presence of adjustment costs: On the one hand, fast emission reductions reduce environmental damage. This is particularly relevant in the case of stock pollutants such as CO₂, because emissions cause damage over several time periods. On the other hand, the cost of emission reductions can always be reduced by extending the time horizon over which emission reductions take place. I show that whereas adjustment costs imply higher emissions along the socially optimal time trajectory, the optimal tax is not reduced in the early periods to give firms time to prepare before the tax is implemented. On the contrary, the socially optimal time trajectory tends to involve large early investments, which gradually declines over time. This contrast with the time profile suggested in the seminal papers by Nordhaus (1991; 1992) on optimal timing of GHG emission reductions, where abatement effort should start low and grows over time. Importantly, and contrary to the DICE model framework used by Nordhaus (1991; 1992), I model technology specific real capital stocks and investment explicitly, such that current investment in relatively clean production technologies decreases the cost of future emission reductions.

I substantiate the analytical findings with a complementary numerical analysis of the U.S. electricity market. The numerical results are in accor-

dance with the analytical findings and shed some light on the numerical magnitudes of the mechanisms examined in the theory section.

The paper features two model extensions. First, I introduce resource scarcity into the model, in which case the well-known dynamics from the literature on the green paradox applies (see, e.g., Sinclair 1992; Sinn, 2008). That is, future taxes decrease the future value of the fossil fuel resource, making it is profitable to move extraction forward in time. Whereas this increases the supply of fossil fuels, the adjustment cost dynamics above reduce the demand for fossil fuels. Hence, it is then theoretically a priori ambiguous whether the market equilibrium will feature increased or decreased fossil fuel consumption, as compared to the case without future taxes. The numerical model provides robust evidence that early emissions will decrease also in the presence of resource scarcity, however. The second model extension relaxes the assumption that the firms' private discount rate equals the social discount rate. It is showed that the standard Pigouvian tax on emissions induces the socially optimal time trajectory in competitive equilibrium only if firms discount the future at a rate equal to the social discount rate. Otherwise, a tax or subsidy on investment is needed. In the numerical example from the U.S electricity market, this involves subsidies to investment in low emission electricity generation capacity, and taxation (or subsidized decommissioning) of investments in emission intensive generation capacity.

The presence of adjustment costs was early recognized; both related to firms' net capital investment decisions (Lucas, 1976; Gould, 1968) and related to changing the number of employees (Holt et al., 1960; Oi, 1962). Capital adjustment costs arise, e.g., if the price of capital increases in the rate of investment. Labor adjustment costs include costs related to hiring, training and layoff. These are all relevant sources for the adjustment costs modelled in the present paper. In the empirical literature, development of models approximating adjustment costs by including lagged dependent variables led to sharp increases in econometric performance (Koyck, 1954; Hall and Jorgenson, 1976). The role of non-convexities and irreversibilities are highlighted by, e.g., Abel and Eberly (1996) and Power (1998). The literature on exhaustible resource extraction with foresighted resource owners is substantial (Hotelling 1931; Heal, 1976), including regulatory issues and the green paradox (Sinclair, 1992; Sinn, 2008).¹ Acemoglu et al. (2012)

¹See also Shapiro (1986), Hamermesh and Pfann (1996), Caballero and Engel (1999),

examines directed technical change in a growth model with environmental constraints. Similar the present paper, albeit for different reasons, they find that delaying regulatory intervention is costly, because it later necessitates a longer transition phase with slow growth.

Despite the early analytical studies of adjustment costs, and the empirical evidence thereof, the literature on regulation in the presence of adjustment costs is scarce. Consequently, our knowledge about how regulatory action should be designed to ease the transition period (where the economy moves from the old equilibrium to the new) is scant. This paper aims to shed some light on that topic.

Section 2.1 is descriptive, examining dynamic effects of suboptimal taxation. Section 2.2 derives the socially optimal time trajectory. Section 2.3 generalizes the results to account for potential resource scarcity and different social and private discount rates. The numerical analysis is in Section 3. Section 4 concludes.

2 Theoretical analysis

Let the vector $\mathbf{x}_t = (x_t^1, x_t^2, \dots, x_t^{\bar{i}})$ denote a representative consumer's consumption bundle of goods $i \in I = \{1, 2, \dots, \bar{i}\}$ in period $t \in T = \{1, 2, \dots, \bar{t}\}$. The associated benefit is given by the increasing and strictly concave utility function $u(\mathbf{x}_t)$. Each good x_t^i is produced by a representative firm (or sector) i . I assume market clearing such that production of x_t^i equals consumption of x_t^i for all $i \in I$ and $t \in T$. The discount factor is given by $\delta \in (0, 1]$ and all derivatives are assumed to be finite.

One interpretation of this model setup is an economy with concave utility from electricity consumption, and where electricity may be derived from \bar{i} energy sources: coal, gas, hydropower, and so forth. To fix thoughts, I will use this as an example throughout the paper.

The investment costs of power generation are essentially capital construction costs and land, including 'regulatory costs' for obtaining siting permits, environmental approvals, and so on. These costs may increase substantially in the presence of economy wide capacity constraints, like limited

Hall (2004), and Cooper and Haltiwanger (2006) about adjustment costs. See Gerlagh (2011), Hoel (2012), Jensen et al. (2015) and Hagem and Storrøsten (2018) for more about the green paradox.

availability of skilled labor or raw materials.² I assume that the investment cost function, $\kappa^i(y_t^i)$, is strictly convex and increasing in investment y_t^i , with minimum at $\kappa^i(0) = 0$. The model framework allows the firm to actively reduce capacity faster than capital depreciation ($y_t^i < 0$).

Production capacity evolves following the state equation:

$$Y_{t+1}^i = \beta Y_t^i + y_t^i, Y_0^i = \bar{Y}^i, \forall i, \forall t, \quad (1)$$

where $\beta \in (0, 1]$ is a capital depreciation factor and \bar{Y}^i is initial capacity (a constant determined by history).

Operating costs for power plants include fuel, labor and maintenance costs. I divide these costs into fixed and variable operating costs. Fixed operating and maintenance costs, denoted $f^i(Y_t^i)$, include, e.g., salaries for facility staff and maintenance that is scheduled on a calendar basis. They do not vary significantly with a plant's electricity generation, but increase in capacity; i.e. we have $\partial f^i(Y_t^i)/\partial Y_t^i \equiv f_{Y_t^i}^i(\cdot) > 0$. The variable operating costs include the cost of consumable materials and maintenance that may be scheduled based on the number of operating hours or start-stop cycles of the plant. These costs are captured by the variable cost function $k^i(x_t^i, Y_t^i)$, with first order derivatives $k_{x_t^i}^i(x_t^i, Y_t^i) > 0$, $k_{Y_t^i}^i(x_t^i, Y_t^i) < 0$, second order derivatives $k_{x_t^i x_t^i}^i(x_t^i, Y_t^i) > 0$, $k_{Y_t^i Y_t^i}^i(x_t^i, Y_t^i) \leq 0$ and cross derivative $k_{x_t^i Y_t^i}^i(x_t^i, Y_t^i) < 0$. Note that variable operating costs decrease in the capacity measure Y_t^i . Total operating costs are then given by:

$$c^i(x_t^i, Y_t^i) = f^i(Y_t^i) + k^i(x_t^i, Y_t^i), \forall i, \forall t. \quad (2)$$

This cost structure implies that operating unit costs, $c^i(x_t^i, Y_t^i)/x_t^i$, have the familiar skewed U-shape with minimum at $c^i(x_t^i, Y_t^i)/x_t^i = c_{x_t^i}^i(x_t^i, Y_t^i)$.

Clearly, the relationship between $f^i(\cdot)$ and $k^i(\cdot)$ may differ markedly across technologies. For example, nuclear power plants feature high fixed costs relative to the variable operating costs, as compared with gas-fueled power plants. The strict convexity of $c^i(\cdot)$ and $\kappa^i(\cdot)$ implies that the 'adjustment costs' associated with any given change in the energy mix \mathbf{x} may be reduced by increasing the number of time periods during which the change

²For example, the modern-day gold rush of oil companies and contractors converging on western Canada's oil-sands markets bogged down as high materials costs and out-stripped labor resources forced project delays and budget overruns around the year 2007; see <http://www.enr.com/articles/29338-oil-sands-boom-extracts-toll-on-costs?v=preview>

occurs. Specifically, the cost of reducing GHG emissions increases with the speed of emission reductions. I let ζ^i denote the emission intensity associated with production of x_t^i .

Let p_t^i denote the consumer price on x_t^i . The competitive representative firm $i \in I$ maximizes the present value of profits over the remaining time horizon solving:

$$\max_{x_t^i, y_t^i} \sum_{t \in T} \delta^{t-1} [(p_t^i - \zeta^i \tau_t) x_t^i - c^i(x_t^i, Y_t^i) - \kappa^i(y_t^i)], \quad \forall t, \quad (3)$$

where τ_t is the tax on one unit of emission. The maximization is subject to equation (1) and has no constraints on the state variable in the last period.³ Note that $p_t^i - \zeta^i \tau_t$ is the producer price after paying the emission tax.

A price-taking representative consumer maximizes net utility solving:

$$\mathbf{x}_t = \arg \max_{\mathbf{x}_t} [u(\mathbf{x}_t) - \mathbf{p}_t \mathbf{x}_t'], \quad \forall t, \quad (4)$$

where $\mathbf{p}_t = (p_t^1, p_t^2, \dots, p_t^{\bar{I}})$. The associated first order condition is $u_{x_t^i}(\mathbf{x}_t^*) = p_t^i$ for all $i \in I$ and $t \in T$.

We have the following result:

Lemma 1. *The competitive equilibrium sequence pair $\{x_t^{i,*}, y_t^{i,*}\}$, solving (3) and (4) subject to equation (1), satisfies:*

$$\begin{aligned} u_{x_t^i}(\mathbf{x}_t^*) - \zeta^i \tau_t &\leq c_{x_t^i}^i(x_t^{i,*}, Y_t^{i,*}), \quad \forall i, \forall t, \\ \lambda_t^{i,*} &= \kappa_y(y_t^{i,*}), \quad \forall i, \forall t, \\ \lambda_t^{i,*} &= -\delta \sum_{r=t+1}^{\bar{t}} (\beta \delta)^{r-t-1} c_{Y_r^i}^i(x_r^{i,*}, Y_r^{i,*}), \quad \forall i, \forall t < \bar{t}, \end{aligned}$$

with $Y_t^{i,*}$ as given by equation (1) and $\lambda_{\bar{t}}^{i,*} = 0$.

Proof. See Appendix A. □

We see from Lemma 1 that production of x_t^i increases in capacity Y_t^i and marginal utility from consumption. The variable λ_t^i is a shadow price representing the present value of the change in future profits caused by a

³That is $Y_{\bar{t}}^i$ is endogenously determined by the intertemporal optimization problem. Note that \bar{t} may be arbitrarily far into the future.

marginal increase in current capacity. In the case where optimal production capacity declines towards a new and lower level (faster than capacity depreciation), higher capacity today induces too high fixed operating costs. Hence, the shadow price λ_t^i is negative. Conversely, λ_t^i is positive if optimal capacity shifts upwards.⁴ Whereas the isolated effect of increased emission taxes is to decrease production (given $\zeta^i > 0$), production of x_t^i in competitive equilibrium may increase in the emission tax τ_t if x_t^i is a low emission good. The reason is that the upward shift in each firm's supply cost functions, caused by the emission tax, increases in the emission intensity of the firm. Therefore, residual demand and equilibrium production of relatively low emission goods increases.

2.1 Dynamic effects of future taxes

Lemma 1 implies that a credible announcement of increased future emission taxes has two key effects in the electricity market:

(a) *Reduced fossil fuel demand from power plants:* Future emission taxes increase the future cost of combusting fossil fuels. The decline in future fossil-fueled electricity generation implies that optimal fossil-fueled power plant capacity will be lower in the future. This reduces the profitability of investment in, e.g., coal fired power plants, and thereby the demand for coal (cf., a lower λ_t^i for emission intensive energy in Lemma 1).

(b) *Increased supply of electricity generated from low emission energy sources:* Future emission taxes imply higher supply costs for fossil-fueled power plants. Hence, low emission electricity generation sources, like renewables or nuclear power, gain a competitive advantage when the tax is implemented. This increases the profitability of investing in low emission electricity generation capacity (cf., a higher λ_t^i for low emission energy in Lemma 1). The associated increase in non-fossil electricity generation capacity reduces the electricity market equilibrium consumption of fossil fuels.

These mechanisms operate with a one period time lag, cf. the capacity state equation (1). Whereas the importance of this lag is negligible if T is measured in short time periods, e.g, months or quarters, it is not unreasonable that it takes some time before the effects of altered investment

⁴We have $\lambda_t^i < (>)0$ if capacity Y^i declines (increases) over time. The first order condition for y_t^i then states that $\kappa_y^i(\cdot) < (>)0$, implying that $y_t^i < (>)0$ because $\kappa^i(\cdot)$ is strictly convex with minimum at $\kappa^i(0) = 0$.

and maintenance decisions influence production and emissions.⁵ The firms investment decisions respond immediately to regulation because of convex capacity investment costs.

More formally, we have the following result:

Proposition 1. *Let the economy be described by the competitive equilibrium in Lemma 1, with $I = \{\text{dirty}, \text{clean}\}$, $\zeta^{\text{dirty}} > 0$, $\zeta^{\text{clean}} = 0$ and interior solutions $x_t^i > 0$ for $i \in I$ and $t \in T$. Let the two goods be substitutes in consumption ($\partial^2 u(\cdot) / \partial x_t^{\text{clean}} \partial x_t^{\text{dirty}} < 0$). Assume the regulator in period $s \in T$ credibly announces increased emission taxes in at least one future time period $u \in T$ ($s + 1 < u \leq \bar{t}$). Then dirty production $x_t^{\text{dirty},*}$ and emissions decrease whereas clean production $x_t^{\text{clean},*}$ increases for all $t > s + 1$.*

Proof. The proposition follows directly from Lemma 1. □

Proposition 1 states that dirty production and emissions decline even before the emission tax increase is implemented (but after it has been announced). The one period lag between the tax announcement and the tax increase is needed for the capacity to adjust (nothing happens in period $s+1$, cf. equation 1). Whereas I have used electricity generation as an example, Proposition 1 is relevant for most regulatory issues, including environmental policies aimed at the lion share of manufactured goods, chemicals, metals and agricultural goods. The important assumption is that the cost structure of the good in question features convex investment costs and capacity constraints (cf., equations 1 and 2). Proposition 1 can be formally generalized to several goods, given appropriate constraints on emission intensities and the cross derivatives of the utility function (see also Section 3). Last, an introduction of a new future tax is a special case of Proposition 1, in which case the tax in period $t < u$ is zero (and positive thereafter).

Besides reducing early emissions, the early tax announcement is important to avoid costly overinvestment in emission intensive production capacity. Moreover, because of convex investment costs, it also reduces the total investment cost of building up increased low-emission production capacity.

⁵Representative lead times for new power plants are, e.g., 2-4 years for renewables, 2-3 years for gas, around 4 years for coal, and 6 years for nuclear (EIA, 2017). Major maintenance decisions and investment in new equipment, like new hydro turbines or replacing old boilers with new steam generators may, however, affect capacity significantly faster.

2.2 The socially optimal tax

In this section I derive the socially optimal tax trajectory. Let $\varsigma = (\zeta^1, \zeta^2, \dots, \zeta^{\bar{i}})$ be a vector of emission intensities associated with production, such that the scalar product $\varsigma \mathbf{x}'_t$ is total emissions at time t (\mathbf{x}'_t is the transpose of \mathbf{x}_t). I assume that the emissions stock evolves following the state equation:

$$S_{t+1} = \alpha S_t + \varsigma \mathbf{x}'_t, \quad S_0 = \bar{S}, \quad \forall t \quad (5)$$

where \bar{S} is a constant determined by history and $\alpha \in [0, 1)$ denotes the stock depreciation factor from one period to the next. Environmental damage from emissions depends on current and historic emission levels and is given by $d(\varsigma \mathbf{x}'_t, S_t)$, where $d(\cdot)$ is weakly convex and increasing in both arguments.⁶ Let welfare W be measured as the present value of utility from consumption net of environmental damages, production costs and investment costs:

$$W = \sum_{t \in T} \delta^{t-1} \left[u(\mathbf{x}_t) - d(\varsigma \mathbf{x}'_t, S_t) - \sum_{i \in I} [c^i(x_t^i, Y_t^i) + \kappa^i(y_t^i)] \right]. \quad (6)$$

The regulator faces a trade-off in the presence of adjustment costs. On the one hand, fast emission reductions reduce environmental damage. On the other hand, the convexity of adjustment costs imply that the cost of emission reductions can always be reduced by extending the time horizon over which emission reductions take place. We have the following result:

Proposition 2. *Let the socially optimal sequence pair $\{x_t^{i,sp}, y_t^{i,sp}\}$ maximize welfare (6) subject to equations (1) to (5). Then, the socially optimal time trajectory can be implemented in competitive equilibrium by the Pigou tax:*

$$\tau_t^{sp} = d_{x_t^i}(\varsigma(\mathbf{x}_t^{sp})', S_t^{sp}) / \varsigma^i + \delta \sum_{r=t+1}^{r=\bar{t}} (\alpha\delta)^{r-t-1} d_S(\varsigma(\mathbf{x}_t^{sp})', S_r^{sp})_r, \quad \forall t < \bar{t},$$

$$\text{with } \tau_{\bar{t}}^{sp} = d_{x_{\bar{t}}^i}(\varsigma(\mathbf{x}_{\bar{t}}^{sp})', S_{\bar{t}}^{sp}) / \varsigma^i.$$

Proof. See Appendix A. □

⁶Whereas stock damage is most relevant for carbon and sulfur dioxides, I allow for associate emissions that causes flow damages. For example, coal plants also emit nitrogen oxides and particulate matter which causes smog.

Note that the expression for τ_t^{sp} is the sum of the marginal current flow damage and the present value of the stream of future marginal stock damages following one additional unit of emissions along the socially optimal time trajectory. This is sometimes referred to as the social cost of carbon in the case of greenhouse gases. The analytical solution for the socially optimal time trajectory $\{x_t^{i,sp}, y_t^{i,sp}\}$ is given in Appendix A. See also the numerical Section 3.2.

We observe that the optimal emissions tax is only indirectly affected by the presence of adjustment costs (via capacity investment which affects environmental damage); i.e., the optimal tax is not reduced during the first years in order to give firms time to adjust before the emission tax is implemented. On the contrary, higher investment costs imply higher emissions and, hence, a higher absolute value shadow price on the emissions stock, which again entails higher optimal emissions taxes (see also Figure 5 in Section 3.2).

Zodrow (1985) examines implementation of an efficiency-enhancing capital taxation reform when the government is concerned about arbitrary reform-induced redistributions and investment causes firms to incur adjustment costs. He finds that optimal reform implementation policy involves immediate but partial enactment only if adjustment costs are convex and sufficiently low. Otherwise, a slower implementation is optimal. Proposition 2 shows that it is optimal to implement the full emission tax at once, and that the implemented tax is higher for larger adjustment costs. Importantly, the framework used to derive Proposition 2 does not account for distributive effects; e.g., because the regulator can use lump-sum transfers to address such concerns. This differs from Zodrow (1985).

2.3 Extensions: Resource scarcity and private discounting

In this Section, I first extend the model to allow for resource scarcity, which is particularly relevant for power plants using fossil fuels such as gas and petroleum. Then I generalize the previous results to the case where the firms may discount the future stronger than the social planner.

2.3.1 Resource scarcity

There is an extensive literature about intertemporal effects induced by future environmental policies in the presence of resource scarcity. In particular,

Sinclair (1992) and Sinn (2008) caution against environmental policies that become more stringent with the passage of time, because such policies will accelerate resource extraction and, thereby, accelerate global warming. The explanation is that increasing taxes decreases the future value of the fossil fuel resource deposits, making it profitable to move extraction forward in time. This argument, often coined ‘the green paradox’, suggests that the potential for environmental policies to curb global warming is limited at best.

Assume that a subset of firms $j \in J = \{\tilde{i} + 1, \tilde{i} + 2, \dots, \bar{i}\}$ use a scarce resource as an input factor in production ($J \subseteq I = \{1, 2, \dots, \tilde{i}, \tilde{i} + 1, \dots, \bar{i}\}$). These firms have variable operating cost function $k^j(x_t^j, Y_t^j) + h^i(X_t^i) x_t^i$, where cumulative production, X_t^j , evolves following the state equation:

$$X_{t+1}^j = X_t^j + x_t^j, X_0^j = \bar{X}^j, \forall j, \forall t. \quad (7)$$

Here \bar{X}^j is a constant and I have normalized units in (7) such that one unit of production requires one unit of resource. Resource scarcity implies that unit operating costs increase in cumulative production; i.e., we have $h_{X_t^j}^j(X^j) > 0$.⁷ Note that firm $j \in J$ is an integrated firm that extracts the fossil fuels needed for electricity generation by itself. The generalized cost function is then:

$$\tilde{c}^i(\cdot) = \begin{cases} f^i(Y_t^i) + k^i(x_t^i, Y_t^i), & \forall i \leq \tilde{i}, \forall t, \\ f^i(Y_t^i) + k^i(x_t^i, Y_t^i) + h^i(X_t^i) x_t^i, & \forall i > \tilde{i}, \forall t, \end{cases} \quad (8)$$

For example, the first $i = 1, 2, \dots, \tilde{i}$ firms may generate electricity using renewables or nuclear energy, whereas the remaining $j = \tilde{i} + 1, \tilde{i} + 2, \dots, \bar{i}$ firms are fossil fueled power plants.

In a model framework consisting of equations (1), (3), (4), (7) and (8), announcement of future emission taxes has one additional key mechanism that operates along with mechanisms (a) and (b) in Section 2.1:

(c) *Increased current supply of fossil fuels:* Future taxes decrease the

⁷Cost that increases with accumulated extraction is frequently used in the resource literature; see, e.g., Heal (1976) and Hanson (1980). As pointed out by Hoel (2012), this specification can approximate the case with a fixed resource stock \bar{X}^j by assuming that $h^j(X_t^j) = \varrho^j$ for $X_t^j < \bar{X}^j$ and $h^j(X_t^j) \rightarrow \infty$ for $X_t^j \geq \bar{X}^j$, where ϱ^j is a fixed unit extraction cost. The framework does not include elements like, e.g., technological progress and new discoveries. I abstract from investment costs that increases in scarcity for simplicity.

future value of the fossil fuel resource. Hence, it is profitable to move extraction forward in time.

This is the well-known (weak) green paradox (see, e.g., Sinclair 1992; Sinn, 2008; Gerlagh, 2011). We therefore have the following:

Corollary 1. *Consider a model defined by equations (1), (3), (4), (5), (6), (7) and (8). Assume the regulator announces a higher future emission tax. Then total emissions in the time interval between the tax announcement and the tax increase will drop if and only if the dynamics caused by adjustment costs (mechanisms (a) and (b) in section 2.1) dominates the scarcity dynamic (mechanism (c) in Section 2.3.1). Further, the optimal tax in Proposition 2 is valid in the combined presence of adjustment costs and resource scarcity.*

Proof. See Appendix A. □

The extension in Corollary 1 is relevant for environmental policies that targets emissions from oil and gas. Whereas the resource scarcity dynamic (c) suggest that exhaustible fossil fuel extraction accelerates following signaling of future environmental policies, the adjustment cost dynamics (a) and (b) explored in the present paper have the opposite effect. From a theoretical point of view, it is therefore a priori unknown whether current emissions increase or decrease following signaling of stringent future climate policy, given that resource exhaustibility and adjustment costs are both present. The numerical results in Section 3 clearly indicate that the capacity constraint mechanisms explored in the present paper strongly dominate the supply side mechanism put forth by the green paradox literature.

Equations (10a) to (10f) in Appendix A show that fossil fueled power plants will become less competitive in the longer run if resource scarcity increases, which gives non-fossil technologies a competitive advantage. This does not only affect future production levels, but also current investment decisions. The theory is ambiguous about whether environmental damage from emissions increase or decrease when fossil fuels become scarcer, however (given that scarcity increases for all fossil fuels). The explanation is that scarcer resources have two opposing effects on emissions from electricity generation: (i) increased emissions caused by substitution away from natural gas to less scarce and more emission intensive coal, and (ii) reduced emissions from the overall decline in fossil fuel reserves. The latter effect

dominates for the simulations graphed in Figure 10 in Appendix B, such that environmental damage decreases in resource scarcity.⁸

2.3.2 Private discounting

Propositions 1 and 2 were derived assuming one single discount factor, δ , applying to all firms and the social planner. This is strong assumption, at least when applied to major environmental challenges like greenhouse gas emissions from power plants and climate change. The Stern Review (Stern, 2007), and the following discussion about appropriate social discount rates in cost-benefit analysis, indicates that social discount rate may be below capital market interest rates, at least in the case of climate change (Weitzman, 2007; Tol and Yohe, 2006). Indeed, the Stern Reviews's conclusions about the need for decisive immediate action hinges on the assumption of a near-zero pure time preference discount rate, which are inconsistent with today's marketplace real interest rates and savings rates (Nordhaus, 2007). In this Section I assume that the firms have private discount factors, ϑ^i , satisfying $0 < \vartheta^i < \delta$.⁹

The firm's maximization problem (3) is then replaced by:

$$\max_{x_t^i, y_t^i} \sum_{t \in T} \vartheta^{i,t-1} [(p_t^i - \zeta^i \tau_t) x_t^i - c^i(x_t^i, Y_t^i) - (\kappa^i(y_t^i) + \theta_t^i y_t^i)], \quad \forall t, \quad (9)$$

where I also assumed that the firms may pay a tax on investment, θ_t^i . This investment tax may take three forms: (i) a standard unit tax on investment if $\theta_t^i > 0$ and $y_t^i > 0$, (ii) a subsidy to decommissioning if $\theta_t^i > 0$ and $y_t^i < 0$, and (iii) a subsidy on investment if $\theta_t^i < 0$. The competitive equilibrium is still given by Lemma 1, but with δ replaced by ϑ^i and $\kappa_{y_t^i}(y_t^i)$ replaced by $\kappa_{y_t^i}^i(y_t^i) + \theta_t^i$ (see the proof of Corollary 2 in Appendix A). We observe that

⁸Acemoglu et al. (2012) finds that use of exhaustible resources in dirty production helps the switch from dirty to clean innovation and production technologies. The situation is less straightforward in the present model framework with numerous goods. E.g., in the power sector, increased scarcity reduces the electricity generation using natural gas (scarce fossil fuel), increases the use of renewables and nuclear, and has ambiguous effects on electricity supply from plants using coal (with or without carbon capture and sequestration).

⁹Goulder and Williams III (2012) argue that we should distinguish between a social-welfare-equivalent discount rate appropriate for determining whether a given policy would augment social welfare and a finance-equivalent discount rate suitable for determining whether the policy would offer a potential Pareto improvement.

Lemma 1 now implies that firms may react differently to future prospects. Specifically, firms with higher discount factors will invest or disinvest more aggressively than their counterparts (who discount the future stronger).

Fowlie (2010) examines the US NOx Budget Program and finds that deregulated plants in restructured electricity markets were less likely to adopt more capital intensive environmental compliance options, as compared to regulated or publicly owned plants. One explanation pointed out by Fowlie (2010) is that public power enjoys significant cost of capital advantages, as compared with private capital. A similar argument is supported by Lemma 1 in the case of private discounting, given that managers of deregulated plants care more about short term profits (to please their shareholders) than their colleagues in the publicly owned plants (i.e., ϑ^i is higher for publicly owned plants). In this case, everything else equal, the absolute value shadow price on capacity would be higher for publicly owned plants. Therefore, emission taxes induce larger investments in public clean production capacity, as compared with privately owned plants.

The emission tax given in Proposition 2 is not able to induce the socially optimal time trajectory alone if $\vartheta^i \neq \delta$. Instead, we now have the following result about optimal taxation:

Corollary 2. *Let the socially optimal sequence pair $\{x_t^{i,sp}, y_t^{i,sp}\}$ maximize welfare (6) subject to equations (1), (2), (4) and (5). Assume the competitive equilibrium solves (1), (2), (4) and (9). Then, the socially optimal time trajectory can be implemented with the following taxes:*

$$\tau_t^{sp} = d_{x_t^i} \left(\varsigma (x_t^{sp})', S_t^{sp} \right) / \varsigma^i + \delta \sum_{r=t+1}^{\bar{t}} (\alpha\delta)^{r-t-1} d_S(\varsigma (x_t^{sp})', S_r^{sp})_r, \quad \forall t < \bar{t},$$

$$\theta_t^{i,sp} = \delta \sum_{r=t+1}^{\bar{t}} (\beta\delta)^{r-t-1} c_{Y_r^i}^i(x_r^{i,sp}, Y_r^{i,sp}) - \vartheta^i \sum_{r=t+1}^{\bar{t}} (\beta(\vartheta^i))^{r-t-1} c_{Y_r^i}^i(x_r^{i,sp}, Y_r^{i,sp}), \quad \forall i, \forall t < \bar{t},$$

with $\tau_{\bar{t}}^{sp} = d_{x_{\bar{t}}^i} \left(\varsigma (x_{\bar{t}}^{sp})', S_{\bar{t}}^{sp} \right) / \varsigma^i$ and $\theta_{\bar{t}}^{i,sp} = 0$.

Proof. See Appendix A. □

Here, τ_t^{sp} is the Pigou tax given in Proposition 2. The optimal investment tax, $\theta_t^{i,sp}$, is the difference between the social planner's and the firm's shadow price on capacity Y_t^i along the socially optimal time trajectory. Note

that the socially optimal investment tax is positive (negative) if production decrease (increase) over time. For example, Corollary 2 may imply decommissioning subsidies to coal plants, whereas investment in renewable energy is subsidized; see Figure 6 in the numerical Section 3.2. Last, Proposition 1 remains valid when equation (3) is replaced with equation (9) (given a not too big disturbing investment tax $|\theta_t^i|$).

3 Numerical analysis: Regulating the U.S. electricity market

According to the Obama Administration (U.S. presidential administration from 2009 to 2017), the United States intended to roughly double its pace of carbon pollution reduction, from 1.2 percent per year on average during the period 2005-2020 to 2.3-2.8 percent per year on average between 2020 and 2025. This target was grounded in analysis of cost-effective carbon pollution reductions achievable under existing law and was intended to keep the U.S. on the pathway to achieve deep economy-wide reductions of 80 percent or more by 2050.¹⁰

In this numerical section, I examine how announcement of future U.S. emission taxes affects current emissions from U.S. electricity generation. I also explore the socially optimal emissions path towards an 80 percent reduction in CO₂ emissions from U.S. electricity production in 2050, as compared with the 2015 emissions level. Emissions from the electric power industry constituted about 35 percent of U.S. energy-related CO₂ emissions in 2016.

The United States generated about 4 thousand terawatt hours of electricity in 2016, of which 30 percent came from coal plants, 34 percent from natural gas and petroleum, 20 percent from nuclear power and 15 percent from renewables.¹¹ This numerical illustration features electricity from these four energy sources, and coal-fired power plants with carbon capture and storage (CCS). Electricity is a homogeneous good and I model electricity generated

¹⁰<https://www.whitehouse.gov/the-press-office/2015/03/31/fact-sheet-us-reports-its-2025-emissions-target-unfccc>.

¹¹Petroleum constituted less than 1%. In the ‘renewables’ category we have the following shares: Hydro = 6.5%, biomass = 1.5%, geothermal = 0.4%, solar = 0.9% and wind = 5.6%. Figures are for net electricity generation. See <https://www.eia.gov/tools/faqs/faq.cfm?id=427&t=3>.

from the different sources as perfect substitutes in consumption. Emission reductions are possible either through abatement (CCS), lower electricity consumption, or through substitution from fossil energy to renewables or nuclear energy. I assume that the U.S. government allows increased nuclear energy production to replace fossil fuels.

The numerical model runs over the time horizon $T = \{2016, 2017, \dots, 2115\}$ and uses the Path solver in GAMS (numerical software) to solve the systems of equations in Appendix A as mixed complementarity problems.¹² I assume a discount rate equal to 4 percent per year. Capital depreciation is set to 0.6 percent per year.¹³ Fuel demand is estimated based on figures from the U.S. Energy Information Administration (EIA), IMF and British Petroleum. Fuel specific investment costs are fetched from EIA. Fuel specific operating costs are calibrated using historic figures from EIA, cost estimates of fossil fueled power plant ramp-up costs (Kumar et al., 2012), and figures for remaining fossil fuel reserves from British Petroleum. Supply of electricity generated from fossil fueled power plants, and gas in particular, is modelled quite flexible, whereas nuclear and renewables must invest in production capacity in order to increase production (with more than a few percent above the 2015 level). GDP enters as an explanatory variable in the econometric modelling of U.S. electricity demand, but is kept constant in the numerical simulations. The reason is that energy efficiency seems to have cancelled out GDP growth in the U.S. electricity market during the last decade or so.¹⁴ See Appendix B for details about the numerical model and data sources, including a test of model fit against history.

3.1 Current effects of future taxes

Consider an emission tax that is announced in the beginning of year 2016. The tax is zero for the period 2016-2024 and 50 USD per ton CO₂ thereafter. Figure 1 graphs the changes in net investment (investment minus capital depreciation) induced by the tax announcement in the period 2016-

¹²See Dirkse and Ferris (1995) and '<http://www.gams.com/>' for information about the Path solver and GAMS.

¹³Nadiri and Prucha (1993) estimates the depreciation rates for physical and R&D capital in the U.S. manufacturing sector to 0.059 and 0.12, respectively.

¹⁴From 2005 to 2016, U.S. GDP has increased with 16%, whereas U.S. electricity consumption has increased with 0.45% (IMF World Economic Outlook Database; EIA November 2017 Monthly Energy Review).

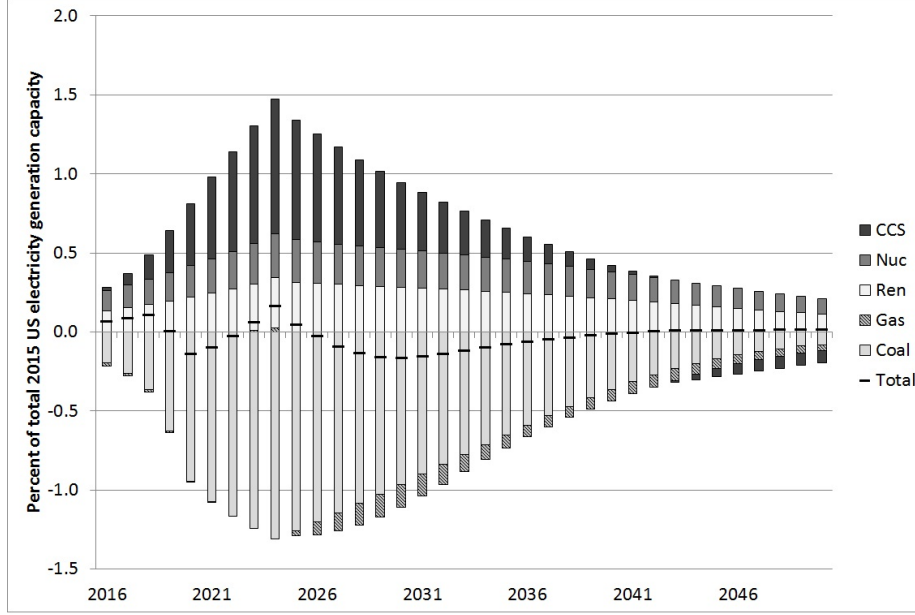


Figure 1: Effects of tax announcement on net investment. Tax minus no-tax simulation values. Tax is 50 USD per ton CO₂ after 2024.

2050. As expected, investment in generation capacity from low emission sources (renewables, nuclear and CCS) increases when the tax is announced. The reason is that future residual demand for electricity from low emission plants will increase when the electricity from coal plants are taxed. In terms of Lemma 1, the emission tax induces a higher future producer price for low emission plants, with an associated higher shadow price on capacity. Furthermore, it is not profitable to invest in coal fired power plants in the face of the future emission tax. Therefore, net investment is negative for coal. The results for gas are less clear. This is as expected from the theory section, because gas is less emission intensive than coal and a scarce resource.¹⁵

Figure 2 shows changes in electricity production and emissions following announcement of the future tax, as compared to the case with no tax. The lower capacity of coal fired power plants, implied by the figures for net investment graphed in Figure 1, causes early production and emission from

¹⁵Leiter et al. (2011) finds a positive but diminishing impact of environmental regulation on investment in an econometric model for European manufacturing industries. This is as expected from Section 2 in the present paper, and in accordance with the larger investment during the early transition phase in Figure 1. The possibility that regulation may induce investment in the period between announcement and implementation is not in discussed in Leiter et al. (2011).

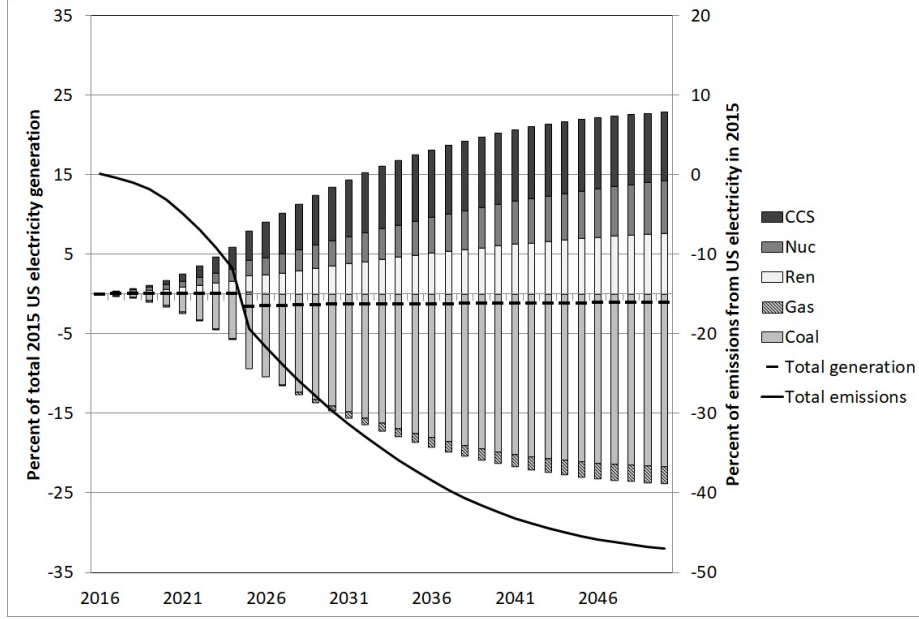


Figure 2: Effects of tax announcement on production and emissions. Production by source (left axis) and total yearly emissions (right axis). Tax minus no-tax simulation values. Tax is 50 USD per ton CO₂ after 2024.

coal to decline. In addition, the increased capacity of low emission power plants crowds out electricity from coal fired power plants, also in the years before the tax is implemented. The black line in Figure 2 shows the associated decline in aggregate yearly emissions. Emissions decline in all periods, except for a minuscule increase in 2016, which occurs because the adjustment cost mechanics operates with a one period time lag (cf., Proposition 1).¹⁶ Overall, the cumulative decline in emissions over the period 2016-2024, i.e. before the tax is implemented, constitutes 39 percent of total emissions in 2015.

Because the firms adjust optimally to the future regulation, Figures 1 and 2 imply that immediate action is optimal even when the emission reduction targets are several years into the future.

How sensitive are the results in Figure 2 with respect to the magnitude of investment costs? In Figure 3, I multiply the model baseline investment costs ($\kappa(\cdot)$) with $\phi \in \{0, 0.05, 0.1, 0.2, 0.4, \dots, 2\}$. Here, $\phi = 0$ is the case

¹⁶The very small increase in emissions in 2016 is caused by resource scarcity dynamics (mechanism (c) in Section 2.3). The capacity constraint mechanisms (a) and (b) from Section 2.1 dominates thereafter.

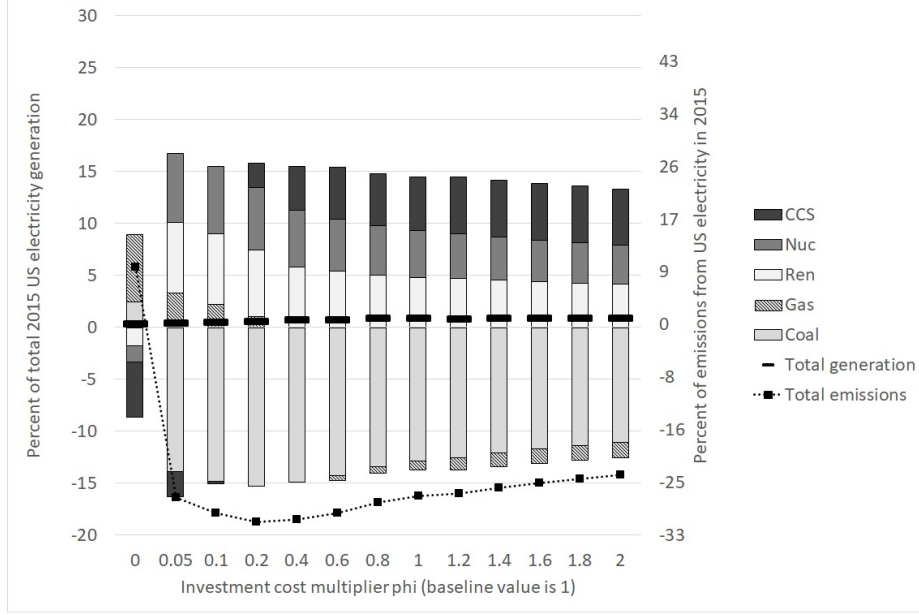


Figure 3: Sensitivity w.r.t. investment costs. Effects of tax announcement on production by source (left axis) and emissions (right axis) summed over the 9 years before the tax is implemented (2016-2024). Tax minus no-tax simulation values.

with costless investment in production capacity, whereas $\phi = 2$ indicates that capacity adjustment costs are doubled. We observe that total emissions from 2016 to 2024 decline unless investment costs are less than 5% of the baseline capital costs calibrated from figures given by the U.S. Energy Administration (EIA, 2016). Note that higher early gas production, caused by announcement of future emission taxes, may crowd out electricity supply from emission intensive coal plants. We also observe that CCS does not enter the market when investment costs are very low. The reason is that CCS has higher operating costs than the non-fossil fuels, making it less profitable to invest in CCS when nuclear and renewable generation capacity expand rapidly. A larger share of CCS reduces the emission reductions graphed in Figure 3, because CCS is more emission intensive than nuclear and renewables.

Increased investment costs have two opposing effects on the change in emissions in Figure 3. On the one hand, larger investment costs implies higher absolute value shadow prices on capacity, which pulls in the direction of a stronger response to the future taxes. On the other hand, higher invest-

ment costs itself imply a weaker response (because it is more expensive to change emission levels). We also observe that total electricity generation in the period 2016 to 2024 increases following the tax announcement. The reason is that the combined effects on production from increased investment in low emission energy sources and a lower scarcity value on fossil fuels (gas in particular) dominates the decrease in production caused by lower capacity of emission intensive energy sources.

A sensitivity analysis was conducted for remaining fossil fuel reserves, GDP growth, discounting, the fuel shares of generation capacity in 2015, and the time lag between tax announcement and tax implementation; see Appendix B. The qualitative results remained unchanged, except for a very small increase in early emissions in the case with a one-period lag only (cf., Proposition 1). Emissions in the time period between tax announcement and tax implementation decreased whenever the time-lag was two years or more. Total emissions over the period 2016-2024 remained lower in the tax simulation (as compared with the no-tax simulation) even when scarcity costs were multiplied with five, and when initial capacity was adjusted such that all electricity in 2015 was generated from gas and petroleum fired power plants. Summa summarum, this suggest that emissions are likely to decline following announcement of future taxes in other energy markets as well (i.e., besides the U.S. market).

3.2 The socially optimal time trajectory

As mentioned in Section 2.3, the regulator faces a trade-off in the presence of investment costs: Whereas faster emission reductions reduce environmental damage, the convexity of investment costs imply that the cost of emission reductions can always be reduced by extending the time horizon over which emission reductions take place. Another implication of this, however, is that economy wide constraints like availability of skilled labor and raw materials may cause emission reductions to take time no matter what. This suggests that immediate action may be necessary to avoid extensive environmental damages in some cases.

Figure 4 graphs net investment along the socially optimal time trajectory in the numerical model, and the (undiscounted) emission tax that implements this trajectory in competitive equilibrium (cf., Proposition 2). We

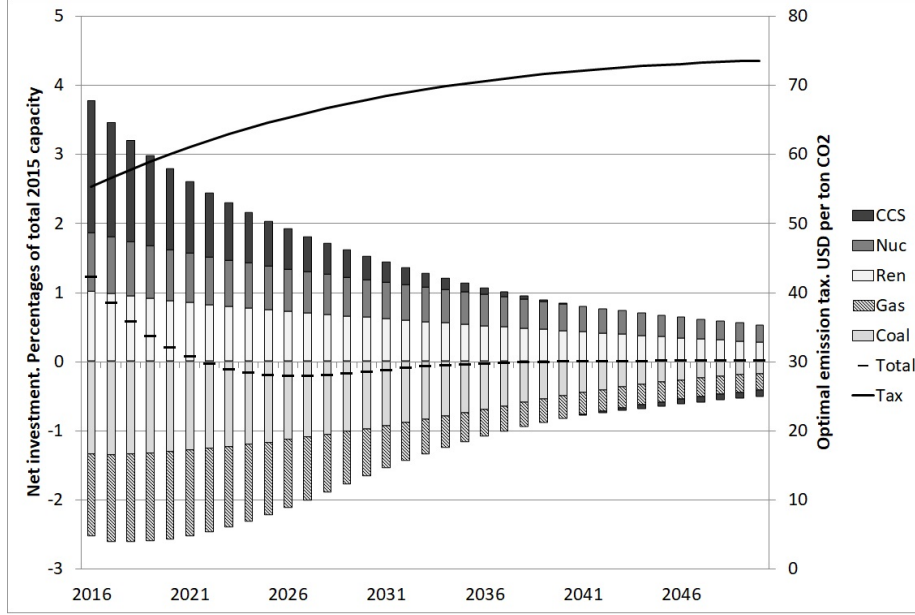


Figure 4: The socially optimal time trajectory: Net investment by source (left axis) and optimal emission taxes (right axis).

observe that the lion share of investment occurs during the first years.¹⁷ Figure 5 graphs optimal taxes and emissions along the socially optimal time trajectory for different assumptions about the magnitude of investment costs. We see that emissions and optimal taxes increase in investment cost. The explanation is that the transition towards a cleaner energy mix slows down when investment costs increase, which again implies higher emissions and larger environmental damage. Hence, the optimal tax increases in investment costs (cf., Proposition 2).

The socially optimal time profile derived in the present paper differs from Nordhaus (1991; 1992). That is, whereas Nordhaus finds that abatement should start low and increase gradually over time, the socially optimal time profile graphed in Figure 4 is characterized by abatement effort that starts out high and then decrease over time. There are two main reasons for this difference: First, contrary to the DICE model employed in Nordhaus (1991; 1992), I explicitly model production capacity, which takes time to change. Second, DICE features exogenous technological growth, which pulls in the direction of delaying abatement effort.¹⁸ Technical change is omitted in the

¹⁷Model experimentation shows that this holds also in the case of a pure flow pollutant.

¹⁸The presence endogenous technological change in the form of learning by doing pulls

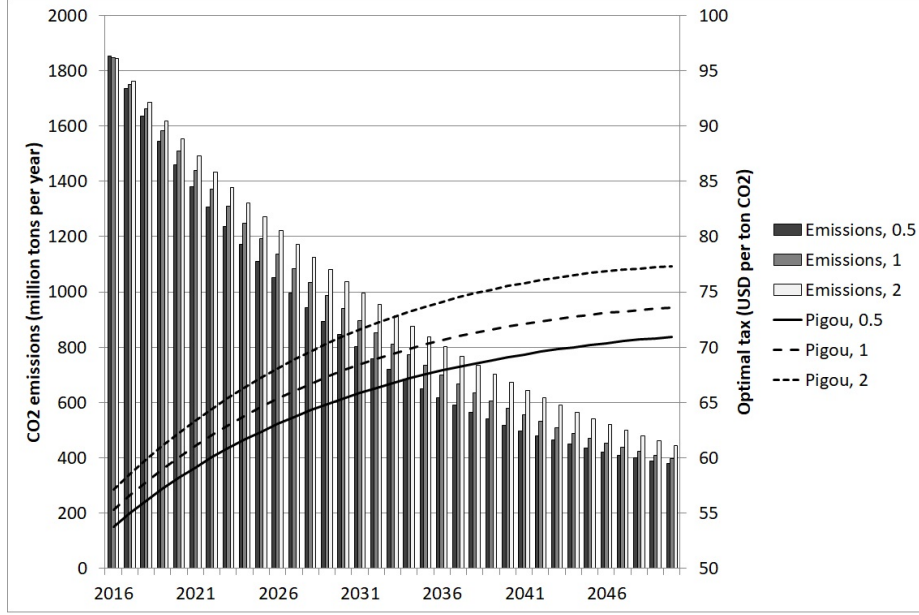


Figure 5: Total emissions and taxes for various assumptions about investment costs. 0.5=halved; 1=baseline (i.e., EIA, 2016); 2=doubled.

main part of present paper, but Appendix B features two different technology scenarios. In the first scenario, renewable investment cost declines by 5 percent each year. This implies, e.g., that renewable investment costs are halved by 2030, and only one tenth of baseline cost in 2060. As expected, exogenous technological change implies a delay in investment. Nevertheless, Figure 10 in Appendix B shows that the optimal trajectory features large investments early on also in this technology optimistic scenario. The second scenario features a clean technology breakthrough. The new technology emerges in 2025, has potential to supply 1500 GWh per year at a marginal supply cost equal to half the 2015 electricity price, and the whole U.S. electricity market at a marginal supply cost equal to the 2015 price. Otherwise, it shares characteristics with the renewable energy sector. As expected, the emergence of revolutionary clean technology has a major impact on the socially optimal time trajectory; see Figure 11 in Appendix B.

Proposition 2 implies that the socially optimal time trajectory can be implemented by a Pigou tax alone only if the private discount factor equals

in the direction of more early abatement; see and Bramoullé and Olson (2005), and Kverndokk and Rosendahl (2007).

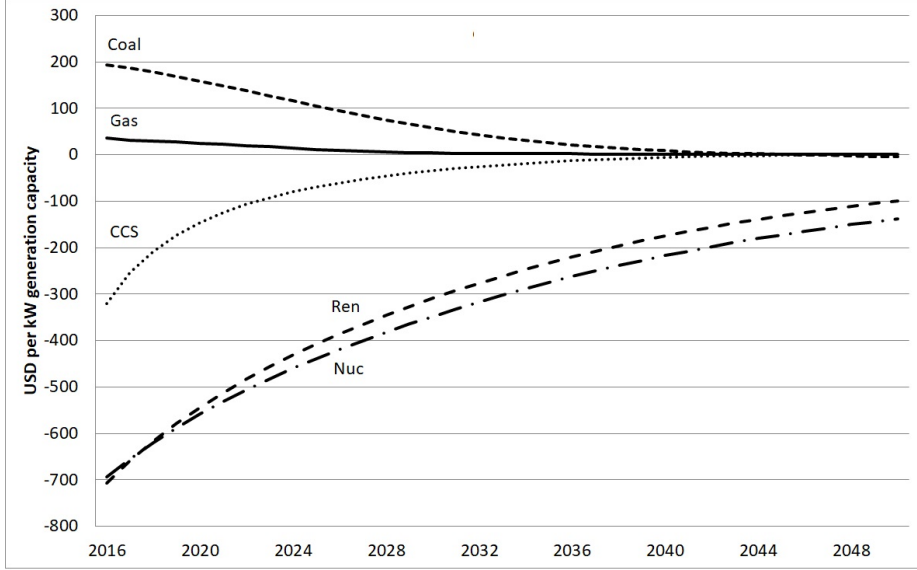


Figure 6: Optimal investment taxes with firm discount factor equal to 0.9 and a social discount factor equal to 0.96. A negative tax indicates a subsidy.

the socially optimal discount factor. Otherwise, the Pigou tax must be supplemented with a tax on investment, cf. Corollary 2. Figure 6 shows the investment taxes/subsidies necessary to induce the socially optimal time trajectory when the private discount factor is 0.9, whereas the social discount factor is 0.96. For comparison, the representative overnight capital costs used in the model calibration are 3636, 6084, 978, 5945 and 2557 USD per kW for coal, CCS, gas, nuclear and renewables, respectively (EIA, 2016).

Interestingly, whereas CCS plays an important part during the transition towards a clean energy mix, the use of CCS declines along the socially optimal time trajectory, as renewable and nuclear capacity increases after a couple of decades. The explanation is that CCS has relatively high operating costs and emissions, as compared to renewables and nuclear power plants. This can be seen from Figure 4, and explains why the subsidies to CCS are relatively low in Figure 6.

4 Conclusion

This paper examined the transition dynamics induced by regulatory action in the presence of convex investment costs. Convex investment costs make it

cost efficient for the firms to start adaptation to anticipated future emission taxes at once. Hence, the economy's production capacity mix immediately begins moving towards cleaner production processes, with associated decreases in emissions. These emission reductions begin before the actual tax increase (given a sufficient time lag between the tax announcement and the tax implementation). The numerical analysis provides convincing evidence that this result remains valid in the presence of resource scarcity. Hence, the analysis suggests that policy makers should strive to keep regulation predictable, with future taxes announced as soon as possible. This helps avoiding erroneous investment decisions and reduces early emissions (the latter contrast with the results from the green paradox literature, see e.g., Sinclair, 1992 and Sinn, 2008).

The results also show that it is inefficient to gradually increase taxation out of concern for the firms' adjustment costs. The optimal tax in a dynamic setting with capacity constraints, investment, and a mix of stock and flow pollutants, is still just the well-known Pigou tax. Note, however, that whereas the principles for the Pigou tax remains unaltered, i.e., the emission tax equals the present value of current and future environmental damages caused by a marginal increase in current emissions along the socially optimal time trajectory, the tax level implied by this principle depends on the magnitude of adjustment costs. Specifically, larger adjustment costs entail a higher emission tax. Moreover, investments towards cleaner production capacity occur early on along the socially optimal time trajectory. All in all, the results suggest that serious environmental damages, even if occurring in the not too far future, are best met with decisive action today. This contrasts with the seminal contributions on optimal timing of GHG emission reductions in Nordhaus (1991; 1992), where abatement effort should start low and grow over time.

Proposition 1 was derived assuming perfect information about future prices. This is a very strong assumption. Nevertheless, the economic rationale is very simple: Expectations about a future emission tax reduces the incentives to maintain and invest in emission intensive production capacity, and increases the incentives to invest in clean alternatives. The associated change in the production capacity mix (i.e., a larger share of clean capacity) causes emissions to decline. As such, the essential assumption is that the tax announcement can induce an increase in expected future emission taxes.

Proposition 2, on the other hand, is clearly not valid without the perfect foresight assumption. That is, a standard Pigou tax (as given in Proposition 2) cannot, in general, induce the socially optimal time trajectory in a dynamic model with investment and imperfect knowledge about future prices.¹⁹

The theory predicts that forward-looking firms will reduce current consumption of goods subject to stringent future regulation. In this respect, it is interesting to observe the current struggle of publicly traded U.S. coal companies.²⁰ Clearly, there are several factors behind this, like slower economic growth, cheap natural gas and current environmental regulation. Nevertheless, it seems reasonable that also bleaker prospects caused by future environmental regulation and increased competition from renewable power partly explain the investors' vanishing interest in coal.²¹

The results in the present paper all hinge on the assumption of strictly convex investment costs. Whereas this is clearly reasonable in the presence of economy wide capacity constraints, or expedited construction of power plants (see the literature on adjustment costs cited in the introduction), there exists plausible scenarios where this convexity is non-existent or negligible. Last, the paper features a stylized model framework and issues like commitment, uncertainty, network externalities and general equilibrium effects are not included in the analysis. It seems reasonable, however, to expect the basic mechanisms explored in the present paper to remain present in a more general setting.

Acknowledgements

Valuable comments by Cathrine Hagem, Knut Einar Rosendahl, Daniel Spiro and Sverre Jensen are gratefully acknowledged. While carrying out

¹⁹A previous version of the present paper examines optimal taxation when the firms' expectations about future prices and taxes are linear combinations of rational and adaptive expectations (Storrøsten, 2017).

²⁰According to Bloomberg (March 17, 2016), the combined market capitalization of U.S. coal miners since 2011 has plunged from over \$70 billion to barely \$6 billion. In the past two years, at least six U.S. coal-mining companies have filed for bankruptcy. Their struggle to find rescue in the financial and capital markets underscores Wall Street's vanishing interest in coal companies (<http://www.bloomberg.com/news/articles/2016-03-16/coal-s-last-man-standing-dragged-to-the-brink-of-bankruptcy>).

²¹The International Energy Administration (IEA) states, referring to the 2015 Paris Climate Conference, that climate policy has emerged as a major driver for the future of coal in large parts of the world (<http://www.iea.org/Textbase/npsum/mtcmr2015sum.pdf>).

this research H.B. Storrøsten has been associated with CREE - Oslo Centre for Research on Environmentally Friendly Energy. The CREE Centre acknowledges financial support from The Research Council of Norway.

References

- Abel, A. B. & Eberly, J. C. (1996), 'Optimal investment with costly reversibility', *Review of Economic Studies* 63, 581.
- Acemoglu, D.; Philippe, A.; Leonardo, B. & David, H. (2012), 'The environment and directed technical change', *The American Economic Review* 102(1), 131-166.
- Bramoullé, Y. & Olson, L. J. (2005), 'Allocation of pollution abatement under learning by doing', *Journal of Public Economics* 89(9), 1935 - 1960.
- Caballero, R. J. & Engel, E. M. R. A. (1999), 'Explaining investment dynamics in US manufacturing: a generalized (S, s) approach', *Econometrica* 67(4), 783-826.
- Cooper, R. W. & Haltiwanger, J. C. (2006), 'On the nature of capital adjustment costs', *The Review of Economic Studies* 73(3), 611-633.
- Dirkse, S. P. & Ferris, M. C. (1995), 'The path solver: a non-monotone stabilization scheme for mixed complementarity problems', *Optimization Methods and Software* 5(2), 123-156.
- EIA (2016), 'Capital cost estimates for utility scale electricity generating plants', Technical report, U.S. Energy Information Administration, Washington DC, U.S.
- EIA (2017), 'Cost and performance characteristics of new generating technologies, Annual Energy Outlook 2017', Technical report, U.S. Energy Information Administration, Washington DC, U.S.
- Fowlie, M. (2010), 'Emissions trading, electricity restructuring, and investment in pollution abatement', *American Economic Review* 100(3), 837-869.
- Gerlagh, R. (2011), 'Too much oil', *CESifo Economic Studies* 57(1), 79-102.
- Gould, J. P. (1968), 'Adjustment costs in the theory of investment of the firm', *The Review of Economic Studies* 35(1), 47-55.
- Goulder, L. H. & Williams III, R. C. (2012), 'The choice of discount rate for climate change policy evaluation', Discussion paper 12-43, Resources for

the Future, Washington DC, U.S.

Hagem, C. & Storrøsten, H. B. (2018), 'Supply- versus demand-side policies in the presence of carbon leakage and the green paradox', forthcoming in *The Scandinavian Journal of Economics*.

Hall, R. E. & Jorgensen, D. (1967), 'Tax policy and investment behaviour', *The American Economic Review* 57(3), 391-414.

Hall, R. E. (2004), 'Measuring factor adjustment costs', *The Quarterly Journal of Economics* 119(3), 899-927.

Hamermesh, D. S. & Pfann, G. A. (1996), 'Adjustment costs in factor demand', *Journal of economic literature* 36, 1264-1292.

Hanson, D. A. (1980), 'Increasing Extraction Costs and Resource Prices: Some Further Results', *The Bell Journal of Economics* 11(1), 335-342.

Heal, G. (1976), 'The relationship between price and extraction cost for a resource with a backstop technology', *The Bell Journal of Economics* 7(2), 371-378.

Hoel, M. (2012), 'Carbon taxes and the green paradox', in R. Hahn & A. Ulph, eds, *Climate change and common sense: Essays in honor of Tom Schelling*, Oxford University Press, chapter 11.

Holt, C. C.; Modigliani, F.; Muth, J. F. & Simon, H. A. (1960), *Planning production, inventories, and work force*, New Jersey, USA: Prentice-Hall.

Hotelling, H. (1931), 'The economics of exhaustible resources', *Journal of Political Economy* 39(2), 137-175.

Jensen, S.; Mohlin, K.; Pittel, K. & Sterner, T. (2015), 'An introduction to the green paradox: The unintended consequences of climate policies', *Review of Environmental Economics and Policy* 9(2), 246-265.

Koyck, L. M. (1954), *Distributed lags and investment analysis*, Amsterdam: North-Holland Publishing Company.

Kverndokk, S. & Rosendahl, K. E. (2007), 'Climate policies and learning by doing: Impacts and timing of technology subsidies', *Resource and Energy Economics* 29, 58-82.

Kumar, N.; Besuner, P.; Lefton, S.; Agan, D. & Hilleman, D. (2012), 'Power Plant Cycling Costs', Technical report, NREL Technical Monitor, Subcontract report NREL/SR-5500-55433 July 2012.

Leiter, A. M.; Parolini, A. & Winner, H. (2011), 'Environmental regulation and investment: Evidence from European industry data', *Ecological Economics* 70(4), 759 - 770.

- Lucas, R. (1976), 'Optimal investment policy and the flexible accelerator', *International Economic Review* 8(1), 78-85.
- Nadiri, M. I. & Prucha, I. R. (1993), 'Estimation of the depreciation rate of physical and R&D capital in the U.S. total manufacturing sector', NBER working paper series. WP. no. 4591, National Bureau of Economic Research.
- Nordhaus, W. D. (1991), 'To slow or not to slow: The economics of the greenhouse effect', *The Economic Journal* 101(407), 920-937.
- Nordhaus, W. D. (1992), 'An optimal Transition path for controlling greenhouse gases', *Science* 258(5086), 1315-1319.
- Nordhaus, W. D. (2007), 'A review of the Stern Review on the economics of climate change', *Journal of Economic Literature* 45(3), 686-702.
- Oi, W. Y. (1962), 'Labor as a quasi-Fixed factor', *Journal of Political Economy* 70(6), 538-538.
- Power, L. (1998), 'The missing link: Technology, investment and productivity', *Review of Economics and Statistics* 80(2), 300-313.
- Shapiro, M. D. (1986), 'The dynamic demand for capital and labor', *The Quarterly Journal of Economics* 101(3), 513-542.
- Sinclair, P. J. N. (1992), 'High does nothing and rising is worse: carbon taxes should keep declining to cut harmful emissions', *The manchester school of economic and social studies* 60(1), 41-52.
- Sinn, H.-W. (2008), 'Public policies against global warming: a supply side approach', *International Tax and Public Finance* 15(4), 360-394.
- Stern, N. H. (2007), *The economics of climate change: the Stern review*, Cambridge, UK, Cambridge University Press.
- Storrøsten, H. B. (2017), 'Regulation in the presence of adjustment costs and resource scarcity: transition dynamics and intertemporal effects', Discussion Papers no. 864; Statistics Norway, oslo, Norway.
- Sydsæter, K.; Hammond, P.; Seierstad, A. & Strøm, A. (2008), *Further mathematics for economic analysis*, Harlow, UK: Pearson Education Limited.
- Tol, R. S. J. & Yohe, G. W. (2006), 'A Review of the Stern Review', *World Economics* 7(4), 233-50.
- Weitzman, M. L. (2007), 'A review of the stern review on the economics of climate change', *Journal of Economic Literature* 45(3), 703.
- Zodrow, G. R. (1985), 'Optimal tax reform in the presence of adjustment

Appendix A: Proofs and derivations

Derivation of Lemma 1. I solve the model with the generalized cost function (8). Lemma 1 then follows as the special case with $\tilde{i} = \bar{i}$. Firm $i \in I$ solves (3) s.t. equation (1). The associated present value Hamiltonian is:

$$H^x = \begin{cases} \delta^{t-1} [(p_t^i - \zeta^i \tau_t) x_t^i - c^i(x_t^i, Y_t^i) - \kappa^i(y_t^i)] + \hat{\lambda}_t^i (\beta Y_t^i + y_t^i), & \forall i \leq \tilde{i}, \forall t \neq \bar{t}, \\ \delta^{\bar{t}-1} [(p_t^i - \zeta^i \tau_t^i) x_t^i - c^i(x_t^i, Y_t^i, X_t^i) - \kappa^i(y_t^i)], & \forall i \leq \tilde{i}, \end{cases}$$

and:

$$H^x = \begin{cases} \delta^{t-1} [(p_t^i - \zeta^i \tau_t) x_t^i - c^i(x_t^i, Y_t^i, X_t^i) - \kappa^i(y_t^i)] + \hat{\lambda}_t^i (\beta Y_t^i + y_t^i) + \hat{\mu}_t^i (X_t^i + x_t^i), & \forall i > \tilde{i}, \forall t \neq \bar{t}, \\ \delta^{\bar{t}-1} [(p_t^i - \zeta^i \tau_t^i) x_t^i - c^i(x_t^i, Y_t^i, X_t^i) - \kappa^i(y_t^i)], & \forall i > \tilde{i}, \end{cases}$$

where $\hat{\lambda}_t^i$ and $\hat{\mu}_t^i$ are the shadow prices on production capacity Y_t^i and cumulative production X_t^i , respectively. The maximum principle for discrete time optimization states that the solution to (3) must satisfy the following necessary conditions for all $i \in I$ (see, e.g., Sydsæter et al., 2008, p. 445):

$$H_{x_t^i}^x = \delta^{t-1} (p_t^{i,*} - \zeta^i \tau_t - c_{x_t^i}^i(z_t^{i,*})) \leq 0, \quad \forall i \leq \tilde{i}, \forall t, \quad (10a)$$

$$H_{x_t^i}^x = \delta^{t-1} (p_t^{i,*} - \zeta^i \tau_t - c_{x_t^i}^i(z_t^{i,*})) + \hat{\mu}_t^i \leq 0, \quad \forall i > \tilde{i}, \forall t, \quad (10b)$$

$$H_{y_t^i}^x = -\delta^{t-1} \kappa_{y_t^i}^i(y_t^{i,*}) + \hat{\lambda}_t^{i,*} = 0 \quad \forall t, \forall i, \quad (10c)$$

$$\hat{\lambda}_{t-1}^{i,*} = H_{Y_t^i}^x = -\delta^{t-1} c_{Y_t^i}^i(z_t^{i,*}) + \beta \hat{\lambda}_t^{i,*}, \quad \forall t \neq \bar{t}, \forall i, \quad (10d)$$

$$\hat{\mu}_{t-1}^{j,*} = H_{X_t^j}^x = -\delta^{t-1} c_{X_t^j}^j(z_t^{j,*}) + \hat{\mu}_t^{j,*}, \quad \forall i > \tilde{i}, \forall t, \quad (10e)$$

$$\hat{\lambda}_t^{i,*} = \hat{\mu}_t^{j,*} = 0, \quad \forall i, \quad (10f)$$

where $z_t^{i,*} = x_t^{i,*}, Y_t^{i,*}$ if $\forall i \leq \tilde{i}$ and $z_t^{i,*} = x_t^{i,*}, Y_t^{i,*}, X_t^{i,*}$ if $\forall i > \tilde{i}$. (10f) is the transversality conditions for the free state variables Y_t^i and X_t^i . The asterisks denote the competitive equilibrium variable values. The assumptions imposed on the cost function $c^i(\cdot)$ ensure that the Hamiltonian is concave along the optimal trajectory for all $t \in T$. Hence, the necessary conditions

above, and the state movement equations (1) and (7), are also sufficient to solve (3) (cf., Arrow's sufficiency theorem).

The solution to $\hat{\lambda}_{t-1}^i = -\delta^{t-1} c_{Y_t^i}^i(\cdot) + \beta \hat{\lambda}_t^i$ in (10d) is $\hat{\lambda}_t^i = \frac{\hat{\lambda}_0^i}{\beta^t} + \sum_{r=t+1}^{r=\bar{t}} \frac{\delta^{r-1}}{\beta^{t-r+1}} c_{Y_r^i}^i(\cdot)$. The transversality condition $\hat{\lambda}_{\bar{t}}^i = 0$ then implies $\hat{\lambda}_0^i = -\sum_{r=1}^{r=\bar{t}} (\beta\delta)^{r-1} c_{Y_r^i}^i(\cdot)$. Inserting in the equation for $\hat{\lambda}_t^i$ above yields $\hat{\lambda}_t^i = -\sum_{r=t+1}^{r=\bar{t}} \delta^{r-1} \beta^{r-t-1} c_{Y_r^i}^i(\cdot)$ ($t < \bar{t}$). The current value shadow price on capacity is then given by:

$$\lambda_t^{i,*} \equiv \frac{\hat{\lambda}_t^{i,*}}{\delta^{t-1}} = -\delta \sum_{r=t+1}^{r=\bar{t}} (\beta\delta)^{r-t-1} c_{Y_r^i}^i(z_r^{i,*}), \quad (11)$$

with $\lambda_{\bar{t}}^{i,*} = 0$.

The solution to $\hat{\mu}_{t-1}^i = -\delta^{t-1} c_{X_r^j}^i(\cdot) + \hat{\mu}_t^i$ in (10e) is $\hat{\mu}_t^i = \hat{\mu}_0^i + \sum_{r=1}^{r=\bar{t}} \delta^{r-1} c_{X_r^j}^i(\cdot)$ for $t < \bar{t}$. The transversality condition $\hat{\mu}_{\bar{t}}^i = 0$ then implies $\hat{\mu}_0^i = -\sum_{r=1}^{r=\bar{t}} \delta^{r-1} c_{X_r^j}^i(\cdot)$. Hence, we have $\hat{\mu}_t^i = -\sum_{r=t+1}^{\bar{t}} \delta^{r-1} c_{X_r^j}^i(\cdot)$ ($t < \bar{t}$). The current value shadow price on cumulative production X_t^i is then given by:

$$\mu_t^{i,*} \equiv \frac{\hat{\mu}_t^{i,*}}{\delta^{t-1}} = -\sum_{r=t+1}^{\bar{t}} \delta^{r-t} c_{X_r^j}^i(x_r^{i,*}, Y_r^{i,*}, X_r^{i,*}) = -\sum_{r=t+1}^{\bar{t}} \delta^{r-t} h_{X_r^i}^i(X_r^{i,*}) x_r^{i,*}, \quad t < \bar{t}, \quad (12)$$

with $\mu_{\bar{t}}^{i,*} = 0$.

Multiply (10a) and (10b) with δ^{t-1} and insert the consumer's first order condition $u_{x_t^i}(\cdot) = p_t^i$ from (4). Lemma 1 then follows from the equations system (10a), (10b), (10c), (11) and (12).

Proof of Proposition 2 and Corollary 1: A benevolent social planner maximizes the present value of welfare solving:

$$W = \max_{\mathbf{x}_t, \mathbf{y}_t} \sum_{t \in T} \delta^{t-1} \left[u(\mathbf{x}_t) - d(\boldsymbol{\varsigma} \mathbf{x}_t', S_t) - \sum_{i \in I} [c^i(\cdot) + \kappa^i(y_t^i)] \right], \quad (13)$$

subject to equations (1) and (5) with no constraints on the state variables in the last period. The maximization is carried out with respect to all $i \in I$. The Hamiltonian is:

$$H^W = \begin{cases} \delta^{t-1} [u(\mathbf{x}_t) - d(\boldsymbol{\varsigma} \mathbf{x}_t', S_t) - \sum_{i \in I} [c^i(x_t^i, Y_t^i) + \kappa^i(y_t^i)]], & \forall i \leq \tilde{i}, \forall t \neq \bar{t} \\ \quad + \sum_{i \in I} [\hat{\lambda}_t^i (\beta Y_t^i + y_t^i) + \hat{\gamma}_t (\alpha S_t + \boldsymbol{\varsigma} \mathbf{x}_t')], & \forall i \leq \tilde{i}, \\ \delta^{\bar{t}-1} [u(\mathbf{x}_{\bar{t}}) - d(\boldsymbol{\varsigma} \mathbf{x}_{\bar{t}}', S_{\bar{t}}) - \sum_{i \in I} [c^i(x_{\bar{t}}^i, Y_{\bar{t}}^i, X_{\bar{t}}^i) + \kappa^i(y_{\bar{t}}^i)]], & \forall i \leq \tilde{i}, \end{cases}$$

and

$$H^W = \begin{cases} \delta^{t-1} [u(\mathbf{x}_t) - d(\boldsymbol{\varsigma}\mathbf{x}'_t, S_t) - \sum_{i \in I} [c^i(x_t^i, Y_t^i, X_t^i) + \kappa^i(y_t^i)]], & \forall i > \tilde{i}, \forall t \neq \bar{t} \\ \quad + \sum_{i \in I} [\hat{\lambda}_t^i (\beta Y_t^i + y_t^i) + \hat{\mu}_t^i (X_t^i + x_t^i) + \hat{\gamma}_t (\alpha S_t + \boldsymbol{\varsigma}\mathbf{x}'_t)], & \forall i > \tilde{i}, \\ \delta^{\bar{t}-1} [u(\mathbf{x}_{\bar{t}}) - d(\boldsymbol{\varsigma}\mathbf{x}'_{\bar{t}}, S_{\bar{t}}) - \sum_{i \in I} [c^i(x_{\bar{t}}^i, Y_{\bar{t}}^i, X_{\bar{t}}^i) + \kappa^i(y_{\bar{t}}^i)]]. & \end{cases}$$

The socially optimal sequence pair $\{\mathbf{x}_t^{sp}, \mathbf{y}_t^{sp}\}$ solving (13) subject to equations (1), (5) and (7) satisfies:

$$\begin{aligned} H_{x_t^i}^W &= \delta^{t-1} (u_{x_t^i}(\mathbf{x}_t^{sp}) - d_{x_t^i}(\boldsymbol{\varsigma}\mathbf{x}_t^{sp'}, S_t^{sp}) - c_{x_t^i}^i(\mathbf{z}_t^{i,sp})) + \zeta^i \hat{\gamma}_t^{i,sp} \leq 0, \quad \forall i \leq \tilde{i}, \forall t \\ H_{x_t^i}^W &= \delta^{t-1} (u_{x_t^i}(\mathbf{x}_t^{sp}) - d_{x_t^i}(\boldsymbol{\varsigma}\mathbf{x}_t^{sp'}, S_t^{sp}) - c_{x_t^i}^i(\mathbf{z}_t^{i,sp})) + \hat{\mu}_t^{i,sp} + \zeta^i \hat{\gamma}_t^{i,sp} \leq 0, \quad \forall i > \tilde{i}, \forall t, \\ H_{y_t^i}^W &= -\delta^{t-1} \kappa_{y_t^i}^i(y_t^{i,sp}) + \hat{\lambda}_t^{i,sp} \leq 0, \quad \forall t, \\ \hat{\lambda}_{t-1}^{i,sp} &= H_{Y_t^i}^{W,sp} = -\delta^{t-1} c_{Y_t^i}^i(\mathbf{z}_t^{i,sp}) + \beta \hat{\lambda}_t^{i,sp}, \quad \forall t \neq \bar{t}, \\ \hat{\mu}_{t-1}^{i,sp} &= H_{X_t^i}^{W,sp} = -\delta^{t-1} c_{X_t^i}^i(\mathbf{z}_t^{i,sp}) + \hat{\mu}_t^{i,sp}, \quad \forall i > \tilde{i}, \forall t \neq \bar{t}, \\ \hat{\gamma}_{t-1}^{sp} &= H_{S_t^i}^{W,sp} = -\delta^{t-1} d_{S_t}(\boldsymbol{\varsigma}\mathbf{x}_t^{sp'}, S_t^{sp}) + \alpha \hat{\gamma}_t^{sp}, \quad \forall t \neq \bar{t}, \\ 0 &= \hat{\lambda}_{\bar{t}}^{i,sp} = \hat{\mu}_{\bar{t}}^{j,sp} = \hat{\gamma}_{\bar{t}}^{sp}, \quad \forall i \in I, \forall j \in J, \end{aligned}$$

where superscript sp denote the variable values along the social planner's optimal time trajectory. The solution to $\hat{\gamma}_{t-1}^{sp} = -\delta^{t-1} d_{S_t}(\boldsymbol{\varsigma}\mathbf{x}'_t, S_t) + \alpha \hat{\gamma}_t$ in the above equation system is $\hat{\gamma}_t = \frac{1}{\alpha^{\bar{t}}} \hat{\gamma}_0 + \sum_{r=1}^{r=t} \frac{\delta^{r-1}}{\alpha^{\bar{t}-r+1}} d_{S_r}(\boldsymbol{\varsigma}\mathbf{x}'_r, S_r)$. The transversality condition $\hat{\gamma}_{\bar{t}} = 0$ then implies $\hat{\gamma}_0 = -\alpha^{\bar{t}} \sum_{r=1}^{r=\bar{t}} \frac{\delta^{r-1}}{\alpha^{\bar{t}-r+1}} d_{S_r}(\boldsymbol{\varsigma}\mathbf{x}'_r, S_r)$. Inserting in the equation for $\hat{\gamma}_t$ above yields $\hat{\gamma}_t = -\sum_{r=t+1}^{\bar{t}} \delta^{r-1} \alpha^{r-t-1} d_{S_r}(\boldsymbol{\varsigma}\mathbf{x}'_r, S_r)$ ($t < \bar{t}$). This is the adjoint related to the emissions stock S_t in the maximization problem (13). The current value shadow price on the emission stock S_t is then:

$$\gamma_t^{sp} \equiv \frac{\hat{\gamma}_t}{\delta^{t-1}} = -\delta \sum_{r=t+1}^{r=\bar{t}} (\alpha \delta)^{r-t-1} d_S(\boldsymbol{\varsigma}\mathbf{x}_r^{sp'}, S_r^{sp}), \quad t < \bar{t}, \quad (14)$$

with $\gamma_{\bar{t}}^{sp} = 0$. Inserting in the equation system with necessary conditions above and rearranging we get:

$$\begin{aligned}
u_{x_t^i}(\mathbf{x}_t^{sp}) &\leq c_{x_t^i}^i(\mathbf{z}_t^{i,sp}) + d_{x_t^i}(\boldsymbol{\varsigma}\mathbf{x}_t^{sp'}, S_t^{sp}) - \zeta^i \gamma_t^{sp}, \quad \forall i \leq \tilde{i}, \forall t, \\
u_{x_t^i}(\mathbf{x}_t^{sp}) &\leq c_{x_t^i}^i(\mathbf{z}_t^{i,sp}) + d_{x_t^i}(\boldsymbol{\varsigma}\mathbf{x}_t^{sp'}, S_t^{sp}) - \zeta^i \gamma_t^{sp} - \hat{\mu}_t^{i,sp}, \quad \forall i > \tilde{i}, \forall t \\
\lambda_t^{i,*} &\leq \kappa_{y_t^i}^i(y_t^{i,sp}), \quad \forall i, \forall t,
\end{aligned}$$

with $\lambda_t^{i,sp}$, $\mu_t^{i,sp}$ and γ_t^{sp} given by equations (11), (12) and (14), and $Y_t^{i,sp}$, S_t^{sp} and $X_t^{i,sp}$ given by equations (1), (5) and (7), respectively. These conditions state the well-known result that current marginal utility from consumption equals the sum of marginal production cost (including the shadow price $\mu_t^{i,sp}$) and marginal environmental damage. It is straightforward to show that the competitive equilibrium equals the socially optimal time trajectory with a tax equal to the social cost of carbon as given in Proposition 2.

Proof of Corollary 2: With private discounting, $0 < \vartheta^i < \delta$, Lemma 1 is replaced with:

$$\begin{aligned}
u_{x_t^i}(\mathbf{x}_t^*) - \zeta^i \tau_t &\leq c_{x_t^i}^i(x_t^{i,*}, Y_t^{i,*}), \quad \forall i, \forall t, \\
-\vartheta^i \sum_{r=t+1}^{\bar{t}} (\beta \vartheta^i)^{r-t-1} c_{Y_r^i}^i(x_t^{i,*}, Y_t^{i,*}) - \theta_t^i &= \kappa_y(y_t^{i,*}), \quad \forall i, \forall t < \bar{t},
\end{aligned}$$

whereas the socially optimal time trajectory solves:

$$\begin{aligned}
u_{x_t^i}(\mathbf{x}_t^{sp}) - d_{x_t^i}(\boldsymbol{\varsigma}\mathbf{x}_t^{sp'}, S_t^{sp}) + \zeta^i \gamma_t^{sp} &\leq c_{x_t^i}^i(x_t^{i,sp}, Y_t^{i,sp}) \quad \forall i \leq \tilde{i}, \forall t, \\
-\delta \sum_{r=t+1}^{\bar{t}} (\beta \delta)^{r-t-1} c_{Y_r^i}^i(x_t^{i,sp}, Y_t^{i,sp}) &\leq \kappa_{y_t^i}^i(y_t^{i,sp}), \quad \forall i, \forall t < \bar{t},
\end{aligned}$$

with γ_t^{sp} given by equation (14) and $\kappa_{y_t^i}^i(y_t^{i,*}) = \kappa_{y_t^i}^i(y_t^{i,sp}) = 0$. These two systems of equations are identical with the taxes given in Corollary 2.

Appendix B: The numerical model

Let the set of electricity generation sources be $I = \{\text{nuclear}, \text{renewables}, \text{coal}, \text{gas}, \text{CCS}\}$, such that x_t^i denotes U.S. electricity produced (and consumed) in year $t \in T$ from energy source $i \in I$. The subset of fossil fuel owners is $J = \{\text{coal}, \text{gas}, \text{CCS}\}$.

The utility function from electricity consumption is given by $u(\mathbf{x}_t) = u_1 \left(\sum_{i \in I} x_t^i \right) - (u_2/2) \left(\sum_{i \in I} x_t^i \right)^2$, where u_1 and u_2 are estimated parameters. The first order condition associated with equation (4) implicitly gives the demand function, given this utility function. I estimate U.S. electricity demand based on yearly figures for U.S. electricity sales to ultimate customers and average yearly prices from the U.S. Energy Information Administration (EIA) over the period 1990-2014, including GDP and the U.S. Henry hub gas price in the regression.²² I let the electricity price in this equation be endogenous and dependent on the U.S. oil price (West Texas Intermediate) and the supply of electricity. The fitted equation system is:

$$ElC = 1806 - 2.992 * ElP + 131.021 * GDP + 13.501 * GasP, (15a)$$

$$ElP = 197.094 - 0.033 * ElC + 0.313 * OilP. (15b)$$

Here electricity consumption (ElC) is measured in TWh, GDP is in trillions of USD (2014), electricity prices (ElP) are in USD (2014) per MWh, gas prices (GasP) are in USD (2014) per million Btu, and oil prices (OilP) are USD (2014) per barrel. All variables are significant at a 5 percent confidence level and the R^2 values are 0.986 and 0.879 for equations (15a) and (15b), respectively. Note the negative sign on electricity consumption (ElC) in (15b). Alternative estimations featuring the real interest rate, wage index and U.S. coal prices give very similar results. One lag Dickey-Fuller unit root tests suggest that U.S. energy consumption and GDP are non-stationary (MacKinnon approximate p-values are 0.37 and 0.73, respectively - the null hypothesis is unit root). However, the one lag Dickey fuller test statistic on the regression residuals is -2.850, implying that we can reject the hypothesis of unit root residuals at a 10 percent confidence level (p-value is 0.0515). This suggests that U.S. GDP and electricity consumption are cointegrated. I derive u_1 and u_2 from equation (15a).

The numerical model employs a cost function that captures the trade-off

²²I use the following data sources: Electricity prices and consumption: Energy Information Administration (EIA) (<http://www.eia.gov/electricity/data.cfm#sales>); U.S. GDP: IMF World Economic Outlook Database (<https://www.imf.org/external/pubs/ft/weo/2015/02/weodata/index.aspx>); gas, oil and coal prices: British Petroleum (<http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html>); wage index: U.S. social security administration (<https://www.ssa.gov/oact/cola/AWI.html>); Interest rate: Federal reserve (<https://www.federalreserve.gov/releases/h15/data.htm>); Inflation: (<http://www.usinflationcalculator.com/inflation/historical-inflation-rates/>).

between fixed and variable operating costs directly:

$$c(x_t^i, Y_t^i) = k(x_t^i) + f(x_t^i, Y_t^i) + h(X_t^i)x_t^i + c_8^i x_t^i.$$

Here, the 'standard' part of $c(\cdot)$ is

$$k^i(x_t^i) = c_1^i x_t^i + \frac{c_2^i}{2} (x_t^i)^2,$$

where c_1^i and c_2^i are fuel specific parameters calibrated such that, for each source, $k_{x_t^i}^i(\cdot)$ equals the average of 1990-2014 real U.S. electricity prices at generation equal to 2015, and doubles at supply equal to total 2015 electricity consumption. $f(\cdot)$ captures the cost of producing at a level that differs from the minimum efficient scale of invested capacity (measured by Y_t^i):

$$f^i(x_t^i, Y_t^i) = g^i(\cdot) (c_4^i + (1 - c_4^i) C^i(\cdot)),$$

with

$$g^i(\cdot) = \frac{c_3^i}{2} \left(\frac{x_t^i - Y_t^i}{0.05Y_0^i + 0.95Y_t^i} \right)^2,$$

$$C^i(\cdot) = \frac{1}{\pi} \sum_{i \in I} \left[\arctan \left(\frac{x_t^i - (0.05Y_0^i + 0.95Y_t^i)}{c_5} \right) + \frac{1}{2} \right].$$

Here c_3^i determines the magnitude of the adjustment costs, c_4^i is the share of adjustment costs that is incurred when production is declining, and c_5 determines the shape of $C^i(\cdot)$. The function $C^i(\cdot)$ is derived using the cumulative Cauchy distribution function. Note that $C^i(\cdot) \in (0, 1)$ and increases steeply from near zero to near 1 around $x_t^i = Y_t^i$, given a low value on c_5 . The use of $0.05Y_0^i + 0.95Y_t^i$ (in place of only Y_t^i) is needed for the numerical model to solve. The parameters c_3^i are calibrated based on historic production and generation capacity figures from EIA, cost estimates of fossil fueled power plant ramp-up costs (including increased capital depreciation from power plant cycling) (Kumar et al., 2012), and fuel specific power plant characteristics. Besides costs related to investment in non-fossil electricity production capacity and shut-down of fossil fueled power plants, relevant adjustment costs may be power grid investments and energy security issues related to renewable energy intermittency. Fossil fuel purchasing costs, c_8^i ,

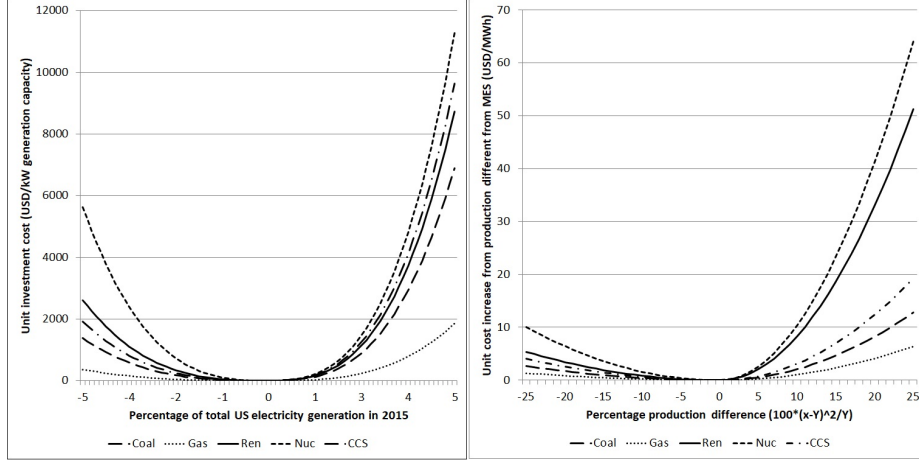


Figure 7: Calibrated investment costs $\kappa(y_t^i)$ on the left and 'adjustment costs' $f^i(x_t^i, Y_t^i)$ on the right.

are assumed constant and equal to figures from EIA for 2015 (c_8^i is zero for renewables and nuclear). Figure 6 graphs the adjustment costs used in the numerical simulations.

I use figures for proved U.S. coal and natural gas reserves from BP Statistics 2016, along with conversion factors and energy content from the Canadian National Energy Board, to derive the resource scarcity unit cost function $h^i(X_t^i) = c_6^i X_t^i$.²³ I assume zero U.S. net imports of coal and gas, and that all U.S. coal and gas resources are available for U.S. electricity production. c_7^i is calibrated such that supply costs of coal and gas doubles when accumulated production X_t^i equals proven reserves (c_7^i is zero for renewables and nuclear). Coal and coal-fired CCS draws from the same fossil fuel resource base.

Fuel specific investment costs are given by:

$$\kappa(y_t^i) = (k_1^i + k_2^i y_t^i) y_t^i (k_3^i + (1 - k_3^i) K^i(\cdot)),$$

with

$$K^i(\cdot) = \frac{1}{\pi} \sum_{i \in I} \left[\arctan\left(\frac{y_t^i}{k_4^i}\right) + \frac{1}{2} \right],$$

where $K(\cdot)$ is derived using the cumulative Cauchy distribution (see com-

²³NEB: <http://www.neb-one.gc.ca/nrg/tl/cnvrsntbl/cnvrsntbl-eng.html>. BP: <http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html>

ments on $C^i(\cdot)$ above). The fuel specific investment costs parameters k_1^i and k_2^i are calibrated using figures from IEA (2016); i.e., overnight capital costs are 3636, 6084, 978, 5945 and 2557 \$/kW for coal, CCS, gas, nuclear and renewables, respectively. Power plants cannot operate all hours during the year, e.g. because of maintenance requirements. Renewables are also dependent on weather conditions. I assume that all power plants can operate 90% of the time, except renewables which has a utilization rate equal to 0.5. I calibrate k_1^i and k_2^i such that unit investment costs equals those given by EIA when the investment level equals average yearly investment during the period 2005-2014. Furthermore, unit investment costs are quadrupled if the investment level is doubled from the average 2005-2014 level. Note that investment tends to occur simultaneously across fuels in the numerical simulation, implying that economy wide capacity constraints may occur. Last, the cost of reducing capacity is 80 percent lower than that of increasing capacity for all fuels except nuclear. The calibrated investment costs are graphed in Figure 7.

Fuel specific emission intensities are based on EIA figures for electricity generation and emissions. I calculate the CCS emissions intensity under the assumption that CCS plants reduce emissions with 90 percent.²⁴ Environmental damage is given by $d(\mathbf{s}\mathbf{x}'_t, S_t) = d_1\mathbf{s}\mathbf{x}'_t + d_2S_t + \frac{d_3}{2}S_t^2$, where d_1 , d_2 and d_3 are calibrated parameters. I set $d_1 = 1$, $d_3 = 0$ and calibrate d_2 such that the Obama Administration's 80% emission reduction target in 2050 (as compared with 2015 emissions) is socially optimal (given the other parameters in the model). See Table 1 for exact parameter values. All prices and costs are measured in 2016 USD in this paper (except for the estimation above). The numerical model solves the systems of equations in Appendix A (social planner and competitive equilibrium) as mixed complementarity problems, given these functional forms and parameter values.

I have tested the model fit by running the model from 2005. Figure 8 shows model projections and historic figures for the period 2005–2015. Production levels, investment levels and electricity prices are endogenous. The simulation features actual values for GDP, coal prices and gas prices (using constant values based on historic averages after 2015). All in all the model does reasonably well, but it struggles with the shale gas revolution.

²⁴CCS has the potential to reduce CO2 emissions from a coal or natural gas-fueled power plant by 90 percent; see <http://www.c2es.org/technology/factsheet/CCS>

Specifically, the supply of coal is too high, whereas the supply of gas is too low in the last years of the sample.

Figure 9 graphs results from a sensitivity analysis w.r.t. the lag between tax announcement and tax implementation (x-axis). We observe that emissions in the time period between tax announcement and tax implementation decrease in all cases; unless the tax is implemented in the next year 2017 (cf. the one-period lag caused by equation 1).

The left diagram in Figure 10 replicates Figure 4 in Section 3.1 in the case of fast technic change, modelled as a 5% yearly reduction in the investment cost parameters k_1^{ren} and k_2^{ren} . We observe that, even though a slower change in the energy mix is optimal in the presence of very fast technological progress, a large share of the investment still occurs in the early years. The right diagram in Figure 10 is a sensitivity analysis w.r.t. resource scarcity (the scarcity parameter c_6^i is 0, 105 and 210 in the simulations denoted with 0, 1 and 2, respectively). Note that the social cost of carbon decreases in resource scarcity.

Figure 11 examines the socially optimal time trajectory in the case of an emerging clean future technology (FuTech). I have omitted CCS and used $T=70$ in this simulation. Further, I allow FuTech to produce 25 GWh per year in the period 2016-2024. This was necessary for the numerical model to solve. The left diagram in Figure 11 replicates Figure 4 in Section 3. The right diagram graphs production levels.

Parameter\fuel	Coal	Gas	Nuclear	renewable	CCS	FuTech
ζ^i	1.011	.3986	.0052	.0052	.101	.0052
c_1^i	46.91	50.56	78.50	89.79	97.42	20.03
c_2^i	.0379	.0352	.0338	.0309	.0250	.0218
c_3^i	750	375	5250	3750	1125	3750
c_4^i	0.2	0.2	0.2	0.2	0.2	0.2
c_6^i	105	105	0	0	105	0
c_7^i	7.7	11.2	0	0	15	0
k_1^i	1.125	.3025	1.839	1.424	1.527	1.424
k_2^i	1.7E-3	4.5E-4	2.2E-3	2.1E-3	1.7E-3	2.4E-2
k_3^i	.25	.25	1	.25	.25	.25
\bar{X}^i	1	1	1	1	1	1
\bar{Y}^i	1341	1241	506	797	25	25
Parameter	α	β	δ	c_5	d_1	d_2
Value	.995	.994	0.96	1.0E-4	1	7.74E-5
Parameter	d_3	k_4	S	u_1	u_2	
Value	0	1.0E-4	1925	1408	.3342	

Table 1: Parameter values in the numerical illustration. c_3^i is measured in thousand USD.

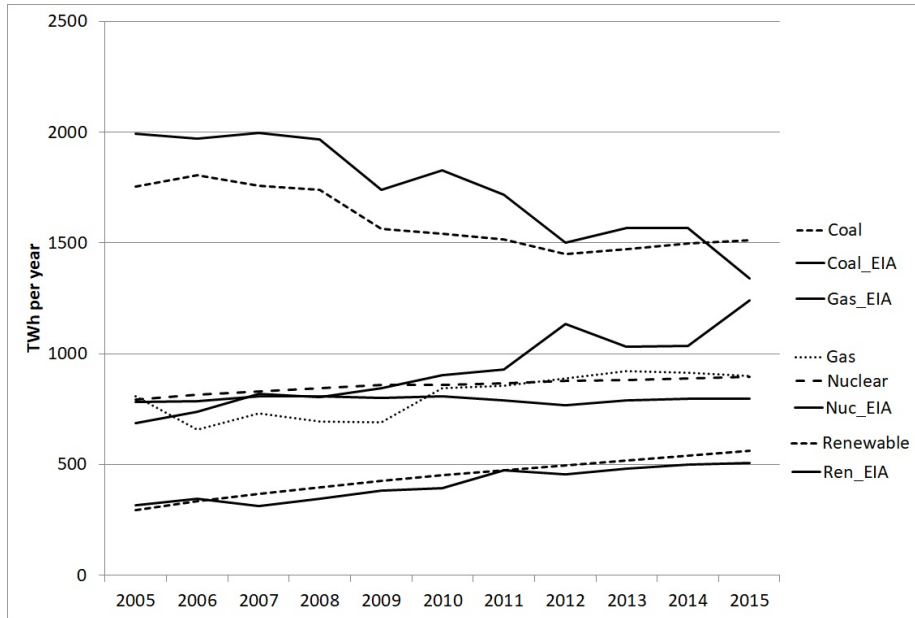


Figure 8: Model fitted equilibrium values (dotted lines) and actual historic values (solid lines).

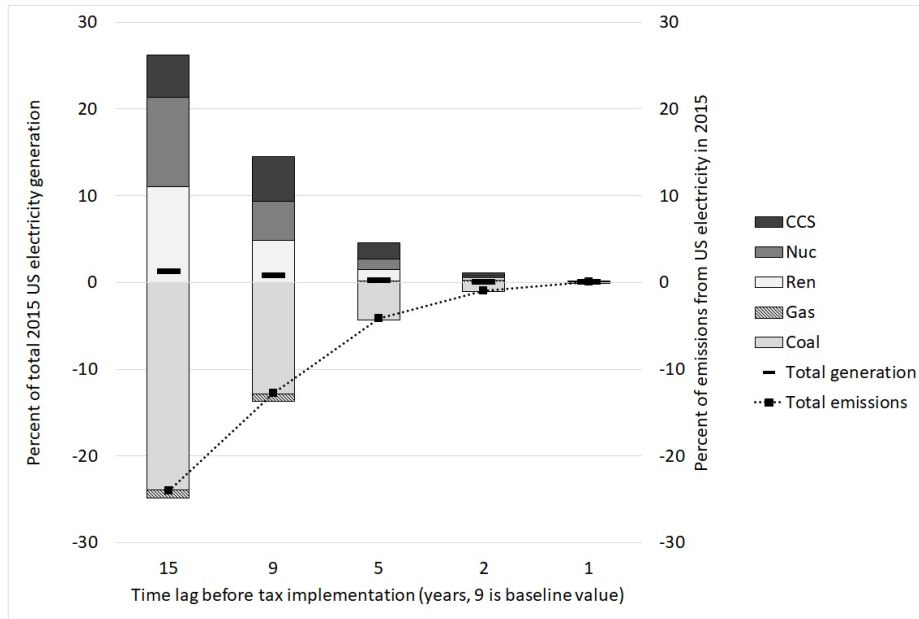


Figure 9: Sensitivity w.r.t. timing of tax implementation. Effects of tax announcement on production by source (left axis) and emissions (right axis) summed over the relevant number of years (given along the x-axis) before the tax is implemented. Tax minus no-tax simulation values.

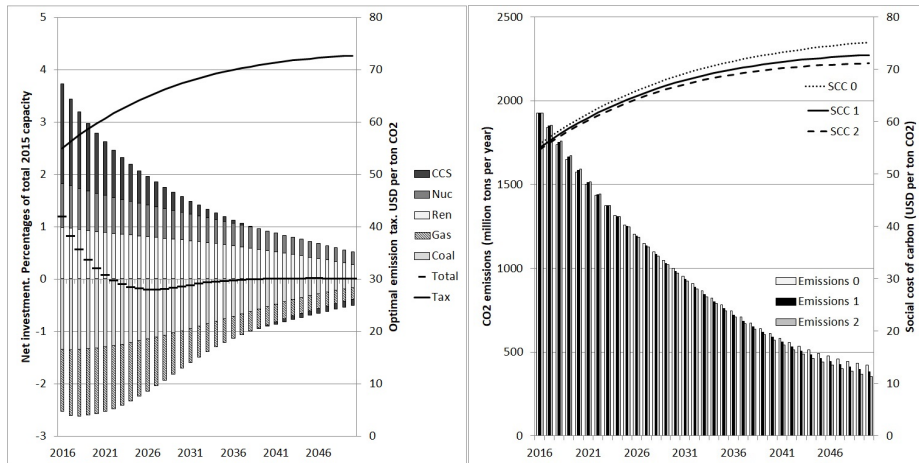


Figure 10: Left diagram: The socially optimal time trajectory with fast renewable technology growth. Net investment by source (left axis) and optimal emission taxes (right axis). Right diagram: Emissions (bars) and the social cost of carbon (lines) when scarcity is zero (0), baseline (1) or doubled (2).

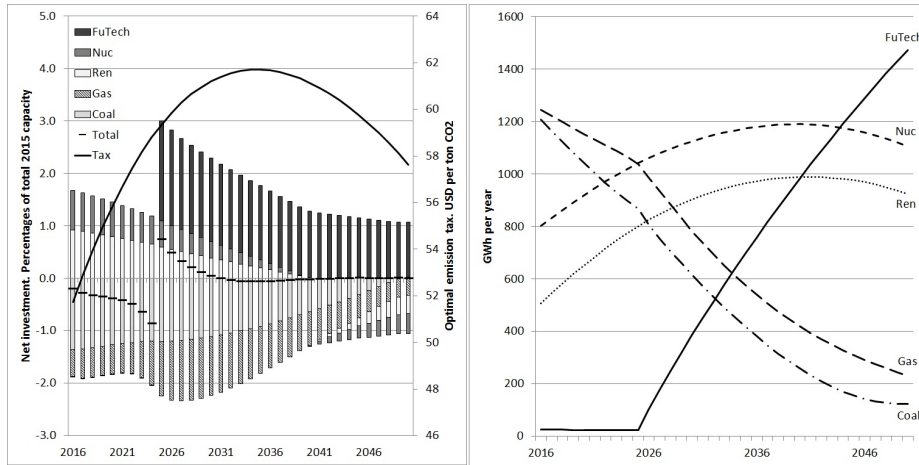


Figure 11: The socially optimal time trajectory with clean technology revolution (FuTech available in 2025). Left diagram: Net investment and optimal tax. Right diagram: Production levels.