Discussion Papers
Statistics Norway
Research department

No. 864
September 2017

Statistics Norway

Statistisk sentralbyrå

Halvor Briseid Storrøsten

Regulation in the presence of adjustment costs and resource scarcity: transition dynamics and intertemporal effects

Discussion Papers No. 864, September 2017
Statistics Norway, Research Department

Halvor Briseid Storrøsten

Regulation in the presence of adjustment costs and resource scarcity: transition dynamics and intertemporal effects

Abstract:
This paper examines regulation in the presence of adjustment costs and resource scarcity, allowing for imperfectly informed firms. I find strong evidence that announcement of future environmental regulation will reduce current emissions in the combined presence of resource scarcity and adjustment costs. This contrasts with the results in the literature on the green paradox. Further, efficient transition towards a low emission economy requires an investment tax on emission intensive production, unless firms have perfect information about the future. Moreover, investments in clean substitutes should first receive a subsidy, but may thereafter be taxed. The optimal tax on production differs from the Pigouvian tax in the case of scarce resources. Last, a uniform tax across heterogeneous agents can induce the socially optimal outcome only if firms have equal expectations about the future.

Keywords: regulation, adjustment cost, imperfect information, exhaustible resources, climate change.

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

Acknowledgements: The financial support of the Norwegian Research Council is highly appreciated. Valuable comments by Knut Einar Rosendahl, Cathrine Hagem, Svenn Jensen and Kjetil Telle are gratefully acknowledged.

Address: Halvor Briseid Storrøsten, Statistics Norway, Research Department. E-mail: hbs@ssb.no

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no/en/forskning/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Sammendrag. Det grønne skiftet: regulering, omstillingsdynamikk og intertemporale effekter

Bedrifter bruker tid til å tilpasse seg nye reguleringer som krever arbeidere med ny kompetanse eller utskiftning av maskiner og bygninger. Dette er relevant både for hvordan regulering virker og for hvordan regulering bør utformes. I dette essayet undersøker jeg temaet med utgangspunkt i virkemiddelbruk og omstilling av elektrisitetssektoren på veien mot lavutslippssamfunnet.

En dynamisk modell viser at annonsering av nye utslippsskatter har tre effekter på utslipp allerede før skattene innføres:

1. *Økt produksjon av begrensede fossile ressurser som olje og gass.* Dette skjer fordi de fremtidige skattene reduserer verdien i å spare resursene for senere produksjon. Ergo er det mer gunstig å utvinne mer nå. Denne mekanismen refereres gjerne til som det «grønne paradokset».
2. *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 brensler.
3. *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.

Mens (1) er en tilbudssideeffekt som drar i retning av økte utslipp, er (2) og (3) etterspørselssideeffekter som bidrar til å redusere utslippene. 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 (2) og (3) dominerer (1), dvs. at utslippene vil falle.

Dersom bedriftenes forventninger om fremtidige priser er adaptive, dvs. at høyere priser i dag gir forventninger om høyere priser fremover, bør utformingen av miljøskatter ta hensyn til dette. Det innebærer at investeringer i forurensende varmekraftverk skattlegges, mens investeringer i relativt rene kraftverk subsidieres. Det viser seg imidlertid at adaptive forventninger og relativt høye priser under omstillingen fra fossil til ikke-fossil energi kan medføre at investeringsbeslutningene i ren produksjonskapasitet baseres på for optimistiske forventninger om fremtidige priser. Dette gir i så fall overkapasitet, for eksempel i form av at for mange fosser legges i rør. For å forhindre dette kan en skattlegge også disse investeringene. Alle investeringsskattene kommer i tillegg til en skatt på utslipp og er kun nødvendige i en overgangsfase.

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 gjerne kreve en kombinasjon av skatter/subsidier på investeringer i en overgangsfase. Skattlegging av utslipp kommer i tillegg.

1 Introduction

A power plant or vehicle may operate for decades before it is obsolete. Consequently, adaptation to new regulatory policies will be sluggish. This is particularly relevant when the regulator wants to induce substantial changes in the economy, like in the case of climate change (see, e.g., IPCC, 2015). In this paper, I examine regulation in the presence of adjustment costs and resource scarcity, allowing for imperfect agent foresight. I focus on climate change and time persistent fossil fuel consumption patterns. Specifically, I assume that firms face convex investment costs, implying that the cost of reducing greenhouse gas (GHG) emissions increases with the speed of emission reductions.

My first research question explores transition dynamics under suboptimal environmental policy; i.e., *does announcement of future emission taxes decrease current emissions in the presence of resource scarcity and adjustment costs*? Anticipated future emission taxes have three effects on early emissions:

(a) Increased current supply of fossil fuels. Future taxes decrease the 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).

(b) Reduced demand for fossil fuels. Future taxes increases the future cost of combusting fossil fuels. This reduces the profitability of investment in, e.g., coal fired power plants, and thereby the demand for coal.

(c) Increased supply of low emission fuel substitutes. Future emission taxes increase future residual demand for low emission energy. This increases the profitability of investment in, e.g., renewable energy and, thereby, the supply of renewable energy to the market. This reduces the consumption of fossil fuels.

Whereas (a) increases the *supply* of fossil fuels, (b) and (c) reduce the *demand* for fossil fuels. Hence, it is a priori ambiguous whether the market equilibrium will feature increased or decreased fossil fuel consumption, as compared to the case without future taxes. Section 3 presents numerical results which suggest that the demand side dynamics (b) and (c) strongly dominate the supply side dynamic (a).

The regulator faces a trade-off in the presence of adjustment costs: On the one hand, fast emission reductions reduce environmental damage from global warming. On the other hand, the cost of emission reductions can always be reduced by extending the time horizon over which emission reductions take place. Furthermore, environmental policy has the dynamic effects (a) to (c) mentioned above. My second research question is: *Should Pigouvian taxes be adjusted in the presence of adjustment costs and resource scarcity*?

Standard Pigouvian taxes induce the socially optimal time trajectory if firms have perfect information about the future. This is not surprising, because the firms then perfectly internalize their future adjustment and resource scarcity costs. But what if the firms are less than perfectly informed about future prices and taxes? To examine this, I let expectations be a linear combination of rational and adaptive expectations.[1] That is, the firms' beliefs about future producer prices (including producer taxes) is a mix between (i) expectations under perfect foresight and (ii) expectations based on a weighted mean of past observations, with less weight on observations further back in time.

It seems reasonable to conjecture that the presence of adaptive expectations induces excess inertia; i.e., that the emission tax must be above the Pigouvian tax level in order to induce efficient transition towards the low emission economy. The rationale is that the firms' investment decisions are based on price expectations that depends on the fossil fuel based economy. Hence, a tax above the Pigouvian level is needed to spur shut-down of emission intensive power plants and investment in low emission energy sources. The answer turns out to be somewhat more complex, however.

Firstly, the optimal time trajectory cannot, in general, be implemented with a tax on emissions alone. The reason is that we have three potential market failures: the negative externality related to emissions, erroneous scarcity rents in the firms' decisions involving extraction of exhaustible resources, and erroneous investment decisions. The first two market failures can be corrected for by a production tax consisting of two elements: a Pigovian tax on emissions and a shadow price element correcting for the

[1]See Muth (1961), Lucas (1987) and Sheffrin (1996) about rational expectations. See Friedman (1957) and Sargent (1999) about adaptive expectations. Chow (1989; 2011) presents econometric evidence in favor of adaptive expectations, as opposed to rational expectations.

erroneous resource rent. The third market failure requires a tax on investment.

Secondly, the presence of adaptive expectations does not necessarily imply excess inertia. The reason is that the clean energy producer price evolves non-monotonically during the transition towards the low emission economy. Clean energy producer prices increase initially, because of higher residual demand for clean energy when the supply of emission intensive energy is taxed. Thereafter, the producer prices decrease as the economy adjusts towards the new equilibrium. Specifically, the production capacity of relatively cheap clean energy replaces the taxed and, hence, more costly emission intensive energy. It follows that adaptive price expectations may be too high during the transition, implying overinvestment in clean energy production capacity and emissions below those of the optimal time trajectory. In the numerical simulations, adaptive expectations induce excess inertia if and only if expectations react sufficiently sluggishly to new information about prices and taxes. One reason for this result is that the clean energy 'producer price spike' that occurs after introduction of emission taxes does not affect the firms' beliefs about future prices as strongly if the adaptive expectation formation process is sluggish. The numerical analysis suggests that the optimal tax on investment in clean energy sources is first negative (subsidy), then positive and slowly declining towards zero. Producer prices on emission intensive energy declines monotonously during the transition towards the low emission economy. The optimal tax on investment in emission intensive production capacity is therefore positive and slowly declining towards zero.

Thirdly, the optimal tax on scarce resources is below marginal environmental damage if producer prices are monotonously decreasing. The reason is that the firms' adaptive producer price expectations are above the actual future prices if prices are declining, implying a too large absolute value scarcity rent. Conversely, the optimal production tax is above marginal environmental damage if producer prices are monotonously increasing. It follows that the optimal production tax on scarce emission intensive resources tend to be below the Pigouvian tax level during the transition towards a low emission economy, because the transition period features declining producer prices for emission intensive goods. After the transition period, however, resource scarcity implies gradually increasing consumer prices (in the absence of technic change). This implies taxation above the Pigouvian level in the long run, unless environmental damage increases sufficiently fast to induce declining producer prices (consumer price minus emission tax). The standard Pigiouvian tax is optimal on production that does not involve exhaustible resources.

Last, a uniform tax does not ensure cost efficiency unless all firms have the same beliefs about the future. The reason is that the perceived shadow prices on production and investment may differ across the firms.

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). There is a substantial literature on exhaustible resources with foresighted resource owners (Hotelling 1931; Heal, 1976), including regulatory issues and the green paradox (Sinclair, 1992; Sinn, 2008).[2]

Section 2 features the theoretical analysis. The numerical Section 3 is included to substantiate selected results and model dynamics. Section 4 concludes.

2 Theoretical analysis

Let the vector $\boldsymbol{x}_t = \left(x_t^1, x_t^2, ..., x_t^{\bar{i}}\right)$ denote a representative consumer's consumption bundle of goods $i \in I = \{1, 2, ..., \bar{i}\}$ in period $t \in T = \left\{1, 2, ..., \bar{t}\right\}$. The associated benefit is given by the increasing and strictly concave utility function $u\left(\boldsymbol{x}_t\right)$. I assume market clearing such that production of x_t^i equals consumption of x_t^i for all $i \in I$ and $t \in T$. One interpretation of this model setup is an economy which uses energy at decreasing returns to scale, and

[2]See also Shapiro (1986), Hamermesh and Pfann (1996), Caballero and Engel (1999), Hall (2004), and Cooper and Haltiwanger (2006) about adjustment costs. See Hoel (2012) and Jensen et al. (2015) for more about the green paradox.

where energy may be derived from $\bar{i}$ sources: coal, gas, hydropower, and so forth. The discount factor is given by $\delta \in (0, 1]$ and all derivatives are assumed to be finite.

Let Y_t^i and X_t^i denote representative firm (or sector) i's time t production capacity and cumulative production (over time), respectively. I assume that the supply (or production) cost of x_t^i is given by:

$$c^i\left(x_t^i, X_t^i, Y_t^i\right) = k^i\left(x_t^i\right) + f^i\left(x_t^i - Y_t^i\right) + h^i\left(X_t^i\right) x_t^i, \ \forall i. \tag{1}$$

Here $k^i(\cdot)$ is a convex and strictly increasing function (standard cost function part), $f^i(\cdot)$ is strictly convex with minimum at $f^i(0) = 0$ (adjustment cost function part), and $h^i\left(X_t^i\right)$ is an increasing function with minimum function value equal to zero (resource scarcity function part); see details below. I assume that supply cost $c^i(\cdot)$ increases in x_t^i such that $\partial c^i(\cdot)/\partial x_t^i \equiv c_x^i(\cdot) \geq 0$ around optimum.[3] Note that the representative firms represent the whole supply chain, including potential resource extraction (mechanisms (a) to (c) in the introduction are internalized by the firm).

The function $f^i(\cdot)$ implies that it is costly to produce at a level that differs from capacity Y_t^i; e.g., because of overtime payments, idle capacity or use of costly reserve capacity. I assume that production capacity evolves following the state equation:

$$Y_{t+1}^i = \beta Y_t^i + y_t^i, \ Y_0^i = \overline{Y}^i, \tag{2}$$

where y_t^i is capacity investment, $\beta \in (0, 1]$ is a capital depreciation factor and $\bar{Y}^i$ is initial capacity (a constant determined by history). The capacity measure Y_t^i may be interpreted as a proxy for minimum efficient scale for production of good x_t^i.

I let investment costs $\kappa^i(y_t^i)$ be a strictly convex function with minimum at $\kappa^i(0) = 0$. The cost of increasing production capacity may consist of building new plants, hiring workers or developing infrastructure. These costs may increase substantially in the presence of economy wide capacity constraints, like limited availability of skilled labor or raw materials.[4] The

[3] The results may be generalized to the case where the cost of producing the different goods depend on each other, given appropriate restrictions on the cross-derivatives.

[4] For 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 outstripped labor resources forced project delays and budget overruns around the year 2007;

model framework allows the firm to actively reduce capacity faster than capital depreciation ($y_t^i < 0$). The adjustment costs in this case represent, e.g., capital costs associated with hastened fossil fueled power plant shut-down or lay-off of workers. The strict convexity of $f^i(\cdot)$ and $\kappa^i(\cdot)$ implies that the adjustment costs associated with any given change in $\boldsymbol{x}$ may be reduced by increasing the number of time periods during which the change occurs. Specifically, the cost of reducing GHG emissions increases with the speed of emission reductions.

The convex and non-decreasing function $h^i\left(X_t^i\right)$ captures potential resource scarcity related to production of x_t^i; i.e., unit cost may increase with cumulative production.[5] The state equation for X_t^i is:

$$X_{t+1}^i = X_t^i + x_t^i,\ X_0^i = \bar{X}^i, \forall i, \tag{3}$$

where $\bar{X}^i$ is a constant.

Let $\boldsymbol{\varsigma} = \left(\zeta^1, \zeta^2, ..., \zeta^{\bar{i}}\right)$ be a vector of emission intensities associated with production (or consumption) of x_t^i. Total emissions at time t is then the scalar product $\boldsymbol{\varsigma x}_t'$ ($\boldsymbol{x}_t'$ is the transpose of $\boldsymbol{x}_t$). I assume that the emissions stock evolves following the state equation:

$$S_{t+1} = \alpha S_t + \boldsymbol{\varsigma x}_t',\ S_0 = \bar{S} \tag{4}$$

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 $d\left(\boldsymbol{\varsigma x}_t', S_t\right)$, where $d(\cdot)$ is weakly convex and increasing in both arguments.[6]

see http://www.enr.com/articles/29338-oil-sands-boom-extracts-toll-on-costs?v=preview

[5]Cost 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 $\widetilde{X}$ by assuming that $h(X) = \varrho$ for $X < \widetilde{X}$ and $h(X) \to \infty$ for $X \geq \widetilde{X}$, where ϱ is a fixed unit extraction cost. The framework does not include elements like, e.g., technological progress and new discoveries.

[6]Whereas 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.

2.1 The socially optimal time trajectory

A benevolent social planner maximizes welfare solving:

$$W = \max_{\boldsymbol{x}_t, \boldsymbol{y}_t} \sum_{t \in T} \delta^{t-1} \left[u\left(\boldsymbol{x}_t\right) - d\left(\varsigma \boldsymbol{x}_t', S_t\right) - \sum_{i \in I} \left[c^i \left(x_t^i, X_t^i, Y_t^i\right) + \kappa^i \left(y_t^i\right) \right] \right], \tag{5}$$

subject to equations (1) to (4) and with no constraints on the state variables in the last period.[7] The maximization is carried out with respect to all $i \in I$. Welfare in (5) is measured as the present value of utility from consumption net of environmental damages, production costs and investment costs. I assume that the social planner has perfect information to derive the socially optimal time trajectory. This allows for comparative analyses of optimal taxes and transition dynamics under various assumptions about the representative firm's knowledge about the future.

Before I present the solution to (5), it is convenient to define the following variables:

$$\lambda_t^{i,z} = -\delta \sum_{r=t+1}^{r=\bar{t}} (\beta\delta)^{r-t-1} f_Y^i \left(x_r^{i,z} - Y_r^{i,z}\right), \ \forall i, \forall t < \bar{t}, \tag{6}$$

$$\mu_t^{i,z} = -\sum_{r=t+1}^{\bar{t}} \delta^{r-t} h_X^i \left(X_r^{i,z}\right) x_r^{i,z}, \ \forall i, \forall t < \bar{t}, \tag{7}$$

$$\gamma_t^z = d_{\varsigma \boldsymbol{x}_t^{z\prime}} \left(\varsigma \boldsymbol{x}_t^{z\prime}, S_t^z\right) + \delta \sum_{r=t+1}^{r=\bar{t}} (\alpha\delta)^{r-t-1} d_S(\varsigma \boldsymbol{x}_t^{z\prime}, S_r)_r, \ \forall t < \bar{t} \tag{8}$$

with $\lambda_{\bar{t}}^{i,z} = \mu_{\bar{t}}^{i,z} = 0$, $\gamma_{\bar{t}}^z = d_{\varsigma \boldsymbol{x}_{\bar{t}}^{z\prime}} \left(\varsigma \boldsymbol{x}_{\bar{t}}^{\prime z}, S_{\bar{t}}^z\right)$ and $z = \{*; t, rat; t, ada\}$. Superscript z indicates three different time trajectories: the socially optimal trajectory $(*)$, the competitive equilibrium time t rational expectations path (t, rat), and the competitive equilibrium time t adaptive expectations path (t, ada) (the paths t, rat and t, ada are derived in Section 2.2 below).

The variable $\lambda_t^{i,*}$ is a shadow price representing the change in future welfare caused by a marginal increase in current capacity $Y_t^{i,*}$. In the case where optimal production capacity declines towards a new and lower level, higher capacity today induces higher future adjustment costs and longer

[7] That is, $X_{\bar{t}}^i$, $Y_{\bar{t}}^i$ and $S_{\bar{t}}$ are endogenously determined by the intertemporal optimization problem.

transition time. Hence, the shadow price $\lambda_t^{i,*}$ is negative. Conversely, $\lambda_t^{i,*}$ is positive if optimal capacity shifts upwards. $\mu_t^{i,*}$ is the shadow price on cumulative production. It is negative in the case of an exhaustible resource, because higher current production then increases future production costs and decreases future welfare. Finally, γ_t^z is the present value of the environmental damage caused by one unit of emissions at time $t \in T$. Note that the expression for γ_t^z is the sum of marginal current flow damage and present value marginal future stock damage. I will henceforth refer to γ_t^z as the social cost of carbon. We have the following result ($\boldsymbol{y}_t = \left(y_t^{clean}, y_t^{dirty}\right)$):

Lemma 1. *The socially optimal sequence pair $\{\boldsymbol{x}_t^*, \boldsymbol{y}_t^*\}$ solving (5) subject to equations (1) to (4) satisfies:*

$$\begin{aligned} \partial u\left(\boldsymbol{x}_t^*\right)/\partial x_t^i &\leq c_x^i(x_t^{i,*}, X_t^{i,*}, Y_t^{i,*}) - \mu_t^{i,*} + \varsigma^i \gamma_t^*, \ \forall i, \forall t, \\ \lambda_t^{i,*} &= \kappa_y^i\left(y_t^{i,*}\right), \ \forall i, \forall t, \end{aligned}$$

with $Y_t^{i,*}$, $X_t^{i,*}$, S_t^*, $\lambda_t^{i,*}$, $\mu_t^{i,*}$ and γ_t^* given by equations (2), (3), (4) (6), (7) and (8), respectively.

Proof. See Appendix A. □

Lemma 1 states the well-known result that current marginal utility from consumption equals the sum of marginal production cost (including the shadow price $\mu_t^{i,*}$) and marginal environmental damage (the emission intensity times the social cost of carbon, $\varsigma^i \gamma_t^*$). We see from Lemma 1 that production $x_t^{i,*}$ tend to be lower if the resource is scarce, or if the environmental damage associated with x_t^i is high. Further, marginal investment cost equals the shadow price on capacity $\lambda_t^{i,*}$ along the socially optimal time trajectory. Otherwise, the social planner could increase present value welfare by changing the investment level.[8] I examine the dynamics of the socially optimal time trajectory in the numerical Section 3.

2.2 The competitive equilibrium time trajectory

Let p_t^i, τ_t^i and φ_t^i denote consumer prices on x_t^i, producer taxes on x_t^i, and investment taxes on y_t^i, respectively (a producer tax is equivalent with an

[8] We 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 $\kappa^i(0) = 0$.

emission tax in this model setup without abatement). A negative tax indicates a subsidy. In competitive equilibrium, a price-taking representative consumer maximizes net utility solving:

$$\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}_t} \left[u\left(\boldsymbol{x}_t\right) - \boldsymbol{p}_t \boldsymbol{x}_t' \right], \; \forall t, \tag{9}$$

where $\boldsymbol{p}_t = \left(x_t^{clean}, x_t^{dirty}\right)$ and with associated first order condition $\partial u\left(\boldsymbol{x}_t\right)/\partial x_t^i = p_t^i$ for all $i \in I$.

The competitive representative firm i maximizes the present value of profits over the remaining time horizon:

$$V_t^i = \max_{x_s^{i,t}, y_s^{i,t}} \sum_{s=t,t+1,\ldots,\bar{t}} \delta^{s-t} \left[\left(p_s^{i,t,e} - \tau_s^{i,t,e}\right) x_s^{i,t} - c^i\left(x_s^{i,t}, X_s^{i,t}, Y_s^{i,t}\right) - \left(\kappa^i\left(y_s^{i,t}\right) + \varphi_s^{i,t,e} y_s^{i,t}\right)\right], \; \forall t \tag{10}$$

subject to equations (1) to (3) and with no constraints on the state variables in the last period. The solution to the dynamic optimization problem (10) depends on the firm's expectations about future prices and taxes. Superscript t, e denotes a period t expectation in (10). We have $p_t^{i,t,e} = p_t^i$, $\tau_t^{i,t,e} = \tau_t^i$ and $\varphi_t^{i,t,e} = \varphi_t^i$, because the firm can observe current prices and taxes. The period t solution to (10) specifies a time trajectory over the remaining periods $s = t, t+1\ldots, \bar{t}$. This trajectory is updated the next period if the firm receives new information about producer prices or investment taxes.

Two prominent approaches for modeling expectations are adaptive expectations and rational expectations.[9] I assume that the firm's decisions which influence the future are based on forecasts that are linear combinations of adaptive expectations and perfectly rational expectations. More precisely, the competitive equilibrium shadow prices $\mu_t^{i,t,c}$ and $\lambda_t^{i,t,c}$ are linear combinations of the shadow prices associated with the trajectories that solves (10) under adaptive and perfectly rational expectations in each period $t \in T$. In the rest of Section 2.2, I first derive these two time trajectories. Then I combine them to model the behavior of a firm with potentially imperfect knowledge about future prices. Last, I derive the competitive equilibrium.

I model adaptive expectations such that the current expectations about

[9] I use the well-known terms 'rational expectations' and 'adaptive expectations' even though the model abstracts from uncertainty .

each future price or tax equals the expectation in the previous period plus an 'error-adjustment' term. This adjustment term raises (lowers) the current expectation if the realized current value turned out to be higher (lower) than expected. The adaptive expectations time t belief about period $t+n$ ($n \in \{1, 2, ..., \bar{t} - t\}$) is given by:

$$\chi_{t+n}^{i,t,ada} = \chi_t^{i,t-1,ada} + \vartheta \left(\chi_t^{i,c} - \chi_t^{i,t-1,ada} \right), \; s.t. \; \chi_1^{i,0,ada} = \bar{\chi}^i, \; \forall i, \tag{11}$$

where $\chi_{t+n}^{i,t,ada} = \left\{ p_{t+n}^{i,t,ada} - \tau_{t+n}^{i,t,ada}, \varphi_{t+n}^{i,t,ada} \right\}$, $\vartheta \in (0, 1]$ is a constant that determines the speed of the error correction adjustment, $\chi_t^{i,c}$ is the realized value in competitive equilibrium (observed in period t), and $\bar{\chi}^i$ is a constant determined by history. The adaptive expectations satisfy a weak form of consistency in the sense that $\chi_{t+n}^{i,t,ada}$ converges towards the true value $\chi_{t+n}^{i,c}$ if all elements in $\left\{ \chi_t^{i,c} \right\}$ remain constant over a sufficiently large time interval. The convergence is only asymptotic unless ϑ equals unity. The adaptive expectations time t control $\left\{ \boldsymbol{x}_s^{t,ada}, \boldsymbol{y}_s^{t,ada} \right\}_{s=t}^{\bar{t}}$ solves (10) subject to equations (1) to (3) and (11). I show in Appendix A that the period t solution is:

$$\begin{aligned} p_s^{i,t,ada} - \tau_s^{i,t,ada} &\leq c_x^i \left(x_s^{i,t,ada}, X_s^{i,t,ada}, Y_s^{i,t,ada} \right) - \mu_s^{i,t,ada}, \; \forall i, \forall t, \\ \lambda_s^{i,t,ada} &\leq \kappa_y^i \left(y_s^{i,t,ada} \right) + \varphi_s^{i,t,ada}, \; \forall i, \forall t, \end{aligned} \tag{12}$$

with $s = t, t+1, \ldots, \bar{t}$, shadow prices $\lambda_s^{i,t,ada}$ and $\mu_s^{i,t,ada}$ given by equations (6) and (7), respectively, and producer prices and taxes given by:

$$p_s^{i,t,ada} - \tau_s^{i,t,ada} = \begin{cases} p_t^i - \tau_t^i & if \; s = t, \\ (1-\vartheta)^t \left(\bar{p}^i - \bar{\tau}^i \right) + \vartheta \sum_{k=1}^t (1-\vartheta)^{t-k} \left(p_k^i - \tau_k^i \right) & if \; s > t, \end{cases} \tag{13}$$

$$\varphi_s^{i,t,ada} = \begin{cases} \varphi_t^i, & if \; s = t, \\ (1-\vartheta)^t \bar{\varphi}^i + \vartheta \sum_{k=1}^t (1-\vartheta)^{t-k} \bar{\varphi}_k^i & if \; s > t, \end{cases}$$

for all $i \in I$. The expected producer prices ($s > t$) in equation (13) solve the difference equation (11). The interpretations of the shadow prices $\mu_s^{i,t,ada}$ and $\lambda_s^{i,t,ada}$ are similar to that of the socially optimal shadow prices $\mu_t^{i,*}$ and $\lambda_t^{i,*}$ given in Section 2.1 above, except that the firm has adaptive expectations about the future and cares about future profits instead of welfare. Note that the equations system (2), (3), (6), (7), (12) and (13) characterizing the adaptive expectations time t trajectory only features current and historic prices and taxes of which the firm has perfect knowledge. Hence, current production and investment is independent of future producer prices along the adaptive expectations trajectory. Investment and, hence, future capacity, increases in historic producer prices along the adaptive expectations path. Note that the adaptive expectations path (12) only differs from the social planner's solution given in Lemma 1 wrt. the shadow prices $\lambda_s^{i,t}$ and $\mu_s^{i,t}$, given the consumer's first order condition to (9) and a Pigouvian tax equal to the social cost of carbon $\tau_t^i = \varsigma^i \gamma_t$.

I now turn to the rational expectations time t control $\left\{\boldsymbol{x}_s^{t,rat}, \boldsymbol{y}_s^{t,rat}\right\}_{s=t}^{\bar{t}}$, which is given by the solution to (10) subject to (1) to (3) under perfect foresight. I show in Appendix A that the period t solution to this optimal control problem is given by:

$$\begin{aligned} p_s^{i,t,rat} - \tau_s^i &\leq c_x^i\left(x_s^{i,t,rat}, X_s^{i,t,rat}, Y_s^{i,t,rat}\right) - \mu_s^{i,t,rat},\ \forall i, \forall t, \\ \lambda_s^{i,t,rat} &\leq \kappa_y^i\left(y_s^{i,t,rat}\right) + \varphi_s^i,\ \forall i, \forall t, \end{aligned} \tag{14}$$

with $s = t, t+1, \ldots, \bar{t}$. The shadow prices $\lambda_s^{i,t,rat}$ and $\mu_s^{i,t,rat}$ in (14) are given by equations (6) and (7), respectively. The rational expectations path induces the socially optimal outcome if the regulator implements a Pigovian tax on emissions (i.e., if $\tau_t^i = \zeta^i \gamma_{t,}$ and given equation 9). Importantly, current production and investment increase in future producer prices along the rational expectations trajectory. The formulation in (14) assumes that taxes are fixed (the perfectly informed social planner does not need to re-optimize).

I assume that all firms always have perfect information about the current state of the system. Hence, the adaptive and rational expectations time t actions, $\left(x_t^{i,t,ada}, y_t^{i,t,ada}\right)$ and $\left(x_t^{i,t,rat}, y_t^{i,t,rat}\right)$, only differ with respect to variables that depends on the future prices and taxes; i.e. the shadow prices $\lambda_t^{i,z}$ and $\mu_t^{i,z}$ (for given initial conditions in period t). I let the representative firms' shadow prices in competitive equilibrium (denoted with superscript c) be linear combinations of the shadow prices under the perfectly rational expectations path and the trajectory associated with adaptive expectations:

$$\lambda_s^{i,t,c} = \psi\lambda_s^{i,t,rat} + (1-\psi)\,\lambda_s^{i,t,ada},\ \forall i, \forall t, \tag{15}$$
$$\mu_s^{i,t,c} = \psi\mu_s^{i,t,rat} + (1-\psi)\,\mu_s^{i,t,ada},\ \forall i, \forall t,$$

with $\psi \in [0,1]$ and $s = t, t+1, \ldots, \bar{t}$. Hence, the representative firms' investment decisions and scarcity considerations are based on a mix between rational and adaptive expectations. Specifically, $\psi = 1$ corresponds to the case of perfect information, whereas $\psi = 0$ amounts to perfectly adaptive expectations. Equation (15) may alternatively be interpreted as modelling an economy with two types of firms within each sector: one with perfectly rational expectations and one with perfectly adaptive expectations. Here, the economy capacity constraints apply to the whole sector producing good x^i, and the parameter ψ determines the relative size of the rational firm type.

We have the following result:

Lemma 2. *The competitive equilibrium sequence pair $\{\boldsymbol{x}_t^c, \boldsymbol{y}_t^c\}$, solving (10) subject to equations (1) to (3) and (9), and with shadow prices being a linear combination of the adjoints associated with perfectly adaptive and perfectly rational expectations as specified in (15), satisfies:*

$$\partial u\left(\boldsymbol{x}_s^c\right)/\partial x_s^i - \tau_s^i \leq c_x^i\left(x_s^{i,c}, X_s^{i,c}, Y_s^{i,c}\right) - \mu_s^{i,t,c},$$
$$\lambda_s^{i,t,c} \leq \kappa_y^i\left(y_s^{i,c}\right) + \varphi_s^i,$$

with $s = t, t+1, \ldots, \bar{t}$; $X_s^{i,c}$and $Y_s^{i,c}$ given by equations (2) and (3), respectively; and where $\mu_s^{i,t,c}$ and $\lambda_s^{i,t,c}$ solves equations *(6), (7) and (12) to (15).*

Proof. See Appendix A. □

The control sequence $\left\{\boldsymbol{x}_s^{t,c}, \boldsymbol{y}_s^{t,c}\right\}_{s=t}^{\bar{t}}$ is a function of the current prices and state variables. This allows the firm to update its trajectory based on the latest available information about the current state of the system; i.e.,$\left\{\boldsymbol{x}_s^{t,c}, \boldsymbol{y}_s^{t,c}\right\}_{s=t}^{\bar{t}}$ is a closed-loop or Markov control. With perfectly rational expectations ($\psi = 1$) there is no need to re-optimize, because the firm perfectly forecasts the future. I will henceforth omit superscript t at $s = t$ to simplify notation when convenient (i.e., we have $\mu_t^{i,t,c} \equiv \mu_t^{i,c}$ and so forth).

The shadow price on capacity $\lambda_t^{i,c}$ depends positively on expected future production levels (cf. equations 6 and 15). Further, the adaptive part of the representative firm's expectations formation causes expectations to lag the actual values whenever $\psi < 1$ and $\{p_t^i - \tau_t^i\}$ is monotonic; i.e., the firm tends to overestimate future producer prices if $\{p_t^i - \tau_t^i\}$ is decreasing, and to underestimate future producer prices if $\{p_t^i - \tau_t^i\}$ is increasing. Therefore, we tend to have $\lambda_t^{i,*} \leq \lambda_t^{i,c}$ if prices are decreasing (and will continue to decrease for a sufficiently long period ahead). Conversely, we tend to have $\lambda_t^{i,*} \geq \lambda_t^{i,c}$ in periods with increasing prices. Because investment depends positively on $\lambda_t^{i,c}$ (cf. Lemma 2), the competive equilibrium features dynamics where periods with high prices and overinvestment are followed by periods with low prices and underinvestment, which again induce high prices and overinvestment, and so forth. The price oscillations decrease over time if taxes τ_s^i and φ_s^i are constant.

The shadow price on cumulative production $\mu_t^{i,z}$ also depends positively on expected future prices (cf., equations 7 and 15). That is, a period with relatively high prices will increase the adaptive price expectations of the firms utilizing a scarce resource as input factor in their production (given $\psi < 1$). The isolated effect of the associated increase in the shadow price is to conserve more of the resource for future use. Conversely, periods with low producer prices tend to feature low absolute value scarcity rents and, hence, stimulate little conservation of the exhaustible resource. I examine the dynamics of the competitive equilibrium trajectory in the numerical Section 3.

2.3 The green paradox revisited

There is an extensive literature about intertemporal effects induced by future environmental policies; see Section 1. In particular, Sinclair (1992) and Sinn (2008) caution against environmental policies that becomes 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, making it profitable to move extraction forward in time (cf., a lower absolute value on $\mu_t^{i,c}$ for fossil fuels in Lemma 2). The green paradox suggests that the potential for environmental policies to curb global warming is limited at best.

In the following, I will use a distinction between a weak and a strong kind of green paradoxes introduced by Gerlagh (2011). The weak green paradox arises when early emissions increase in response to future environmental policies, because fossil fuel owners accelerate production when the future value of the resource stock drops. The strong green paradox arises when the intertemporal adjustment of the resource owners increases not only early emissions, but also the present value of total environmental damages.

We have the following result:

Proposition 1. *Let the economy be described by the competitive equilibrium in Lemma 2, with $\psi > 0$, $I = \{clean, dirty\}$, $\zeta^{dirty} > 0$, $\zeta^{clean} = 0$ and $h_X^{clean}(\cdot) = 0$. Assume interior solutions such that $x_t^i > 0$ for $i \in I$ and $t \in T$. Let the two goods be substitutes in consumption ($\partial^2 u(\cdot) / \partial x_t^{clean} \partial x_t^{dirty} < 0$). Consider a credible announcement at $t = 1$ about future emission taxes $\tau_u^{dirty} > 0$ for all $u = \{v, v+1, ...\bar{t}\}$ ($v \in T \setminus \{1\}$). We then have the following:*

(a) *Let $f^i(\cdot) = 0$ ($\forall x_t^i$) and $h_X^{dirty}(\cdot) > 0$. Then $x_t^{dirty,c}$ increases whereas $x_t^{clean,c}$ decreases for all $t < u$ (weak green paradox).*

(b) *Let $f^i(\cdot) \neq 0$ ($\forall x_t^i \neq Y_t^i$) and $h_X^{dirty}(\cdot) = 0$. Then $x_t^{dirty,c}$ decreases whereas $x_t^{clean,c}$ increases for all $t > 1$ (opposite of green paradox).*

(c) *Let $f^i(\cdot) \neq 0$ ($\forall x_t^i \neq Y_t^i$) and $h_X^{dirty}(\cdot) > 0$. Then either a) or b) above occurs, depending on which of the opposing mechanisms that dominates. In either case, there is an increase in $x_1^{dirty,c}$ and a decrease in $x_1^{clean,c}$.*

Proof. The proposition follows from Lemma 2. □

An increase in $x_t^{dirty,c}$ in Proposition 1 implies increased emissions. Announcement of future taxes has no effect on current emissions in the case of purely adaptive expectations ($\psi = 0$).

Part *a*) in Proposition 1 is the well-known (weak) green paradox (Sinclair, 1992; Sinn, 2008). This holds in the case with resource scarcity and no adjustment costs.

Part *b*) in Proposition 1 is the case with adjustment costs and no resource scarcity. In this case, announcing a future emission tax will reduce early dirty production and emissions. The explanation is that the shadow price on the dirty good production capacity decreases if a future tax is announced. The associated lower investment reduces production capacity and, thereby, dirty good production and emissions (cf., a lower $\lambda_t^{dirty,c}$ in Lemma 2). The exception in the first period occurs because capacity operates with a one period lag in this model (cf., equation 2). 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 decisions influence production and emissions. Furthermore, Proposition 1 *b*) states that anticipation of future emission taxes will affect production of the clean good. The reason is that the clean good firm knows that residual demand for the clean good will increase when the future tax on the dirty good is implemented. Hence, the value of the clean good capacity stock increases. The firm starts investing in the first period because of convex capacity investment costs (cf., a higher $\lambda_t^{clean,c}$ in Lemma 2). Part *b*) in Proposition 1 is relevant for the majority of environmental policies. Examples include non-fossil energy sources and perhaps coal (which exists in abundance), or dirty versus clean manufactured or agricultural goods. We observe that the dynamic effects in part *b)* help to decrease early emissions even before the tax is implemented (but after it has been announced).

Part *c*) in Proposition 1 adds the two mechanisms in parts *a*) and *b*) together. The case with both adjustment costs and resource scarcity is relevant for environmental policies that targets emissions from oil and gas.[10] Whereas the resource scarcity dynamics put forward by the green paradox suggest that exhaustible fossil fuel extraction accelerates following signaling of future environmental policies, the adjustment cost dynamics 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 agents are foresighted and that resource exhaustibility and adjustment costs are present. The numerical results in Section 3 suggest that the capacity constraint mechanisms explored in the present paper strongly dominate the supply side mechanism put forth by the green paradox literature.

So far, we have focused on intertemporal effects induced by suboptimal taxation. I now turn to the issue of optimal taxation in the presence of

[10]There are generally significant capital investment costs related to extraction of oil and gas (see, e.g., IEA 2016, p. 144-160).

adjustment costs and potentially imperfect firm foresight.

2.4 Optimal taxation in the presence of adjustment costs and imperfect knowledge about future prices

In Section 2.2 we found that the competitive equilibrium is characterized by alternating periods of over- and underinvestment, and too little or too much conservation of scarce resources. The optimal taxes must account for these dynamics and correct for the negative environmental externality. We have the following result:

Proposition 2. *Let the economy be described by Lemma 1 and Lemma 2. Then the socially optimal time trajectory can be implemented in competitive equilibrium by the following sequence of taxes:*

$$\tau_t^{i,*} = \varsigma^i \gamma_t + \mu_t^{i,c} - \mu_t^{i,*}, \ \forall i, \forall t,$$
$$\varphi_t^{i,*} = \lambda_t^{i,c} - \lambda_t^{i,*}, \ \forall i, \forall t < \bar{t},$$

Proof. The proposition follows from Lemma 1 and Lemma 2. Note that $\mu_t^{i,c} = \mu_t^{i,rat} = \mu_t^{i,*}$ and $\lambda_t^{i,c} = \lambda_t^{i,rat} = \lambda_t^{i,*}$ when $\psi = 1$ and $\tau_t^{i,*} = \varsigma^i \gamma_t$. □

Proposition 2 implies that a standard Pigouvian tax on emissions induces the socially optimal time trajectory if and only if expectations are perfectly rational, i.e. we have $\tau_t^{i,*} = \varsigma^i \gamma_t$ and $\varphi_t^{i,*} = 0$ iff $\psi = 1$. If expectations are partly adaptive, however, optimal taxation involves two additional 'shadow price elements'. These elements are an investment tax, $\lambda_t^{i,c} - \lambda_t^{i,*}$, and a resource conservation tax, $\mu_t^{i,c} - \mu_t^{i,*}$. In the following, I will examine these elements one by one.[11]

To simplify the discussion of the investment tax, consider the case with no resource scarcity ($h_X^i(\cdot) = 0$), partly adaptive expectations ($\psi < 1$) and adjustment costs $f^i(\cdot) \neq 0$ ($\forall x_t^i \neq Y_t^i$). In this case, optimal policy must correct for two sources of market failure: the negative environmental externality and erroneous investment decisions caused by adaptive expectations. Whereas the Pigouvian tax corrects for the negative environmental externality, the investment tax is still required to correct for the erroneous

[11] In a setting with abatement, such that emissions and production are decoupled, optimal taxation would involve an investment tax $\varphi_t^{i,*} = \lambda_t^{i,c} - \lambda_t^{i,*}$, a Pigouvian tax on emissions $\tau_t^{i,pig*} = \varsigma^i \gamma_t$ and a conservation tax on production $\tau_t^{i,cons*} = \mu_t^{i,c} - \mu_t^{i,*}$.

expectations. Interestingly, it turns out that over- or underinvestment creates negative externalities that extends beyond those related to the investing firms' own emissions. To see this, consider the case with two substitute goods $I = \{clean, dirty\}$. Assume that the dirty producer overinvests in capacity. This will have three consequences. Firstly, the dirty good producer loses profits due excess investment costs and production capacity. Secondly, residual demand for the clean good decreases, implying lower production and less investment. Thirdly, high investment in dirty good production capacity reduces current prices, and hence the clean producer's adaptive expectations about future prices. Therefore, investment in dirty good production capacity reduces the expected profitability from investment in clean good production capacity and, thereby, reduces future clean good capacity and production. This increases future emissions. By the same reasoning, underinvestment in dirty good production capacity induces excess clean good production capacity. Note that the absolute value of λ_t^i will be relatively high during a transition phase, and relatively low when production is stable. Therefore, the taxes or subsidies on investment $\varphi_t^{i,*}$ will be small unless the economy is in a transition phase.[12]

Regarding the conservation tax element, $\mu_t^{i,c} - \mu_t^{i,*}$, Proposition 2 implies that the optimal tax $\tau_t^{i,*}$ is below (above) marginal environmental damage if producer prices are monotonously decreasing (increasing) and $\psi < 1$. The reason is that the firms' adaptive producer price expectations are above (below) the actual future prices if prices are declining (increasing), implying a too large (small) absolute value scarcity rent. This implies that the conservation tax element tends to be negative during the transition period, because the producer prices are steadily declining as the economy adjusts towards the low emission economy. After the transition has been completed, however, consumer prices start to increase due resource scarcity. If this causes producer prices to increase, the conservation tax element will be positive in the long run.

[12]The trajectories in Lemma 1 and 2 may have stationary states. If so, these are characterized by: (i) no use of exhaustible resources; (ii) no stock pollution ($\alpha = 0$ or $d_S(\cdot) = 0$) or that the quantity of pollution added to the emissions stock in each period is equal to the amount that depreciates (so that net stock accumulation is zero) and; (iii) that the firm's expectations are correct. The expectations only approach the true value asymptotically unless $\psi = 1$ and/or $\vartheta = 1$ (cf. equation 11). If existing, the stationary states along the competitive and socially optimal trajectories are equal if and only if the optimal taxes given by Proposition 2 is implemented.

Corollary 1 examines the relationship between the long run conservation tax element and the environmental damage function:

Corollary 1. *Assume one good $I = \{carbon\}$ with negligible adjustment costs ($\lambda_t^{carbon} \approx 0$), resource scarcity ($\mu_t^{carbon} < 0$) and (at least partly) adaptive expectations ($\psi < 1$). Then we have:*

(a) *Production of x_t^{carbon} should be taxed above marginal environmental damage ($\tau_t^{carbon,*} > \varsigma^{carbon}\gamma_t$) if marginal environmental damage from emissions is non-increasing over time ($\gamma_{t+1} \leq \gamma_t$, $\forall t$).*

(b) *Production of x_t^{carbon} should be taxed below marginal environmental damage ($\tau_t^{carbon,*} < \varsigma^{carbon}\gamma_t$) if marginal environmental damage from emissions increases sufficiently fast over time ($\gamma_{t+1} \gg \gamma_t$, $\forall t$).*

Proof. The corollary follows directly from Proposition 2. □

The first part of Corollary 1 is valid in the case of no environmental damage ($d(\cdot) = 0$), or if environmental damage increases sufficiently slowly over time. The explanation is that the firm producing *carbon* does not fully internalize the increase in the consumers' marginal utility from consumption of x^{carbon} induced by the aggravating future resource scarcity. Hence, the regulator increases the current tax to conserve some of the exhaustible resource for future use. Note that the producer price must increase over time for the first part in Corollary 1 to apply.[13] The second part of Corollary 1 relates to the case where producer prices $p_t^i - \tau_t^i$ decrease over time. This occurs if marginal environmental damage from emissions increases sufficiently fast over time. In this case, the producer's adaptive price expectations are too high, implying a too high resource rent and, consequently, excess conservation of the resource. Intuitively, it is better to extract a larger share of the resource today if future marginal environmental damage is high.

Proposition 2 implies that the optimal taxes $\tau_t^{i,*}$ and $\varphi_t^{i,*}$ depend on the firms' beliefs about future prices. How does this affect cost efficiency if we momentarily relax the assumption about one single representative firm producing good i? We then have the following result:

[13] It is straightforward to show that part (a) in Corollary 1 may be negated by technic change; i.e. the conservation tax element is negative if technic change dominates resource scarcity in the long run (such that producer prices declines).

Corollary 2. *Let there be $j \in J = \{1, 2, ..., \bar{j}\}$ = firms producing good $i \in I$. Then the optimal tax is uniform if and only if $\psi^{i,j} = 1$ for all firms, or if $\psi^{i,j}$ and $\vartheta^{i,j}$ are both equal across firms ($\forall i \in I$, $\forall j \in J$ and $\forall t \neq \bar{t}$).*

Proof. Different expectations about the future implies different shadow prices (cf. Lemma 2). The corollary then follows from Proposition 2. □

Corollary 2 states that a uniform tax across heterogeneous agents can induce the socially optimal path only if firms have equal expectations about the future. Indeed, even 'static cost efficiency', in the sense of equal marginal supply costs across firms in period $t < \bar{t}$, cannot be guaranteed. The intuition is straightforward: the forward-looking firm's current production end investment decisions depends on the shadow price on production and investment, which again depends on the firms' expectations about the future. If the shadow prices differ across firms, equal tax rates cannot ensure that marginal supply costs are equalized across firms (cf., Lemma 1). Corollary 1 implies that market based regulatory instruments, like uniform taxes or tradable quantity regulation, cannot ensure cost efficiency unless expectations about the future are equal across all firms.

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

According to the White House, the United States intends 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 is grounded in analysis of cost-effective carbon pollution reductions achievable under existing law and will keep the U.S. on the pathway to achieve deep economy-wide reductions of 80 percent or more by 2050.[14]

In this numerical illustration, I consider an 80 percent reduction in CO2 emissions generated by U.S. electricity production in 2050, as compared with the 2015 emissions level. Below I give a brief non-technical description of the numerical model. See Appendix B for further details.

[14]https://www.whitehouse.gov/the-press-office/2015/03/31/fact-sheet-us-reports-its-2025-emissions-target-unfccc.

The United States generated about 4 thousand terawatt hours of electricity in 2015, of which 33 percent came from coal plants, 34 percent from natural gas and petroleum, 20 percent from nuclear power and 13 percent from renewables.[15] This numerical illustration features electricity from these four energy sources, and carbon capture and storage (CCS). Electricity is a homogeneous good and I model electricity generated 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 a discount rate equal to 3 percent per year. The numerical model runs over the time horizon $t = \{2016, 2017, ..., 2215\}$ and uses the Path solver in GAMS (numerical software) to solve the systems of equations given in Lemma 1 and Lemma 2 as mixed complementarity problems.[16] I let $\bar{t}$ be large such that the model approximates the infinite horizon solution for the reported results.

I estimate electricity demand using yearly figures for U.S. electricity consumption and prices from the U.S. Energy Information Administration (EIA) over the period 1990-2014, including U.S. GDP figures from the IMF World Economic Outlook Database and fossil fuel prices from BP Statistical Review of World Energy 2016 in the regression. I further assume that $k^i(\cdot)$ is convex such that, for each source, $k^i_x(\cdot)$ 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. Capital depreciation is set to 0.6 percent per year.[17] Regarding investment costs $\kappa^i(\cdot)$, I first calibrate fuel specific adjustment cost factors based on investment cost figures from IEA.[18] Then I scale average investment costs such that it equals 25 times the 2015 electricity price when all energy sources increases capacity with 10 percent of total 2015 electricity generation. The adjustment cost function $f^i(\cdot)$ is calibrated such that marginal supply costs increaes with 105

[15]Petroleum constituted only 1%, whereas natural gas generated 33%. In the 'renewables' category we have the following shares: Hydro = 6%, biomass = 1.6%, geothermal = 0.4%, solar = 0.6% and wind = 4.7%. Figures are for net electricity generation. See https://www.eia.gov/tools/faqs/faq.cfm?id=427&t=3.

[16]See Dirkse and Ferris (1995) and 'http://www.gams.com/' for information about the Path solver and GAMS.

[17]Nadiri 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.

[18]See 'https://www.eia.gov/analysis/studies/powerplants/capitalcost/'

USD per MWh (the U.S. average electricity price in 2015) when production is 20 percent above minimum efficient scale. It is significantly less costly to decrease capacity and to produce below minimum efficient scale than to build new plants and produce above capacity in the model. 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. 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, and that half of the CCS plants combust coal and the other half combust gas.[19] The only modeled environmental damage externality is related to CO2 emissions, and I assume that the U.S. government allows increased nuclear energy production to replace fossil fuels. Last, the adaptive expectations producer price forecast in 2015 is calculated based on historic electrity prices and equation (11). See Appendix B for further details on the numerical model.

3.1 Current effects of future taxes

In this section, I examine how the U.S. electricity market responds following announcement of future CO2 taxes. The tax is announced in the beginning of year 2016. It is zero for the period 2016-2024 and 50 USD per ton CO2 thereafter. I assume perfectly rational expectations in Section 3.1 ($\psi = 1$).[20]

Figure 1 graphs the changes in net investment (investment minus capital depreciation) induced by the tax announcement in the period 2016-2050. As expected, investment in generation capacity from low emission sources (renewables, nuclear and CCS) increase 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 2, 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.

Figure 2 shows changes in electricity production and emissions follow-

[19] CCS 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

[20] $\psi < 1$ does not alter the qualitative results in Section 3.1.

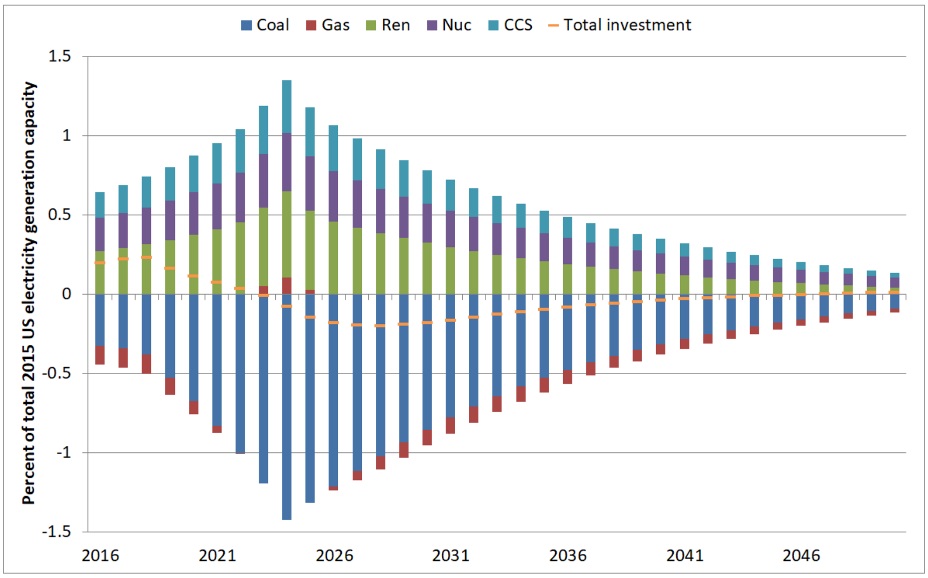

Figure 1: Effects of tax announcement on net investment. Tax minus no tax simulation values. Tax is 50 USD per ton CO2 after 2024.

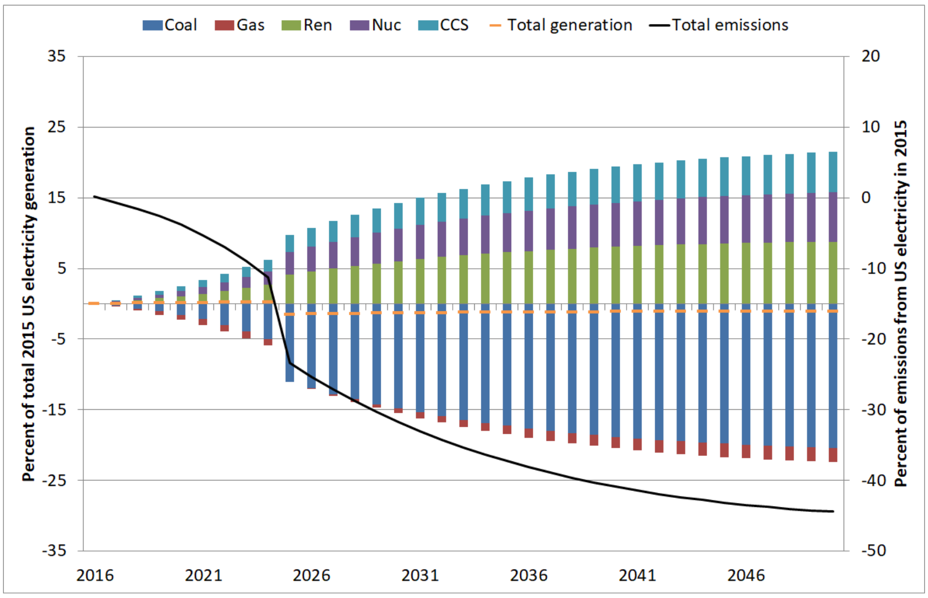

Figure 2: Effects of tax announcement on production and emissions. Production by source (left axis) and total emissions (right axis). Tax minus no tax simulation values. Tax is 50 USD per ton CO2 after 2024.

ing announcement of the future tax, as compared to the case with no tax. The lower capacity of coal fired power plants implied by the net investment graphed in Figure 1 causes early production and emission from coal to decline. In addition, the increased capacity of low emission power plants crowds out electricity from coal and gas fired power plants, also in the years before the tax is implemented. The black line in Figure 2 shows the associated decline in aggregate emissions. Reduced electricity generation from coal account for 92 percent of the total emission reduction over the time interval 2016-2050. 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). Overall, the cumulative decline in emissions over the period 2016-2024, i.e. before the tax is implemented, constitutes 41 percent of total emissions in 2015. Even emissions from gas and petroleum over the period 2016-2024 decline with 81 million tons of CO2 in the tax simulation, as compared with the no tax simulation run.

How sensitive are the result in Figure 2 with respect to the magnitude of adjustment costs? In Figure 3, I multiply the model baseline adjustment costs $(\kappa(\cdot))$ and $(f(\cdot))$ with $\phi \in \{0, 0.2, ..., 2\}$. Here, $\phi = 0$ is the case with costless adjustment, whereas $\phi = 2$ indicates that adjustment costs are doubled. We observe that the numerical model reproduces the weak green paradox if and only if adjustment costs are very low.[21] Interestingly, even emissions from gas and petroleum decreases in all sensitivity cases, except for $\phi = 0$. Note that larger adjustment costs have two opposing effects on the change in emissions in Figure 3. One the one hand, higher adjustment costs implies larger absolute value shadow prices on capacity, which pulls in the direction of a stronger response to the future taxes. On the other hand, higher adjustment costs in itself imply a weaker response (because it is more expensive to change emission levels).

A sensitivity analysis with respect to resource scarcity did not yield the weak green paradox; see Fig. 12 in Appendix B. Specifically, total emissions over the period 2016-2024 remained significantly lower in the tax simulation (as compared with the no-tax simulation) even when scarcity costs were multiplied with five. U.S. electricity generation from gas and petroleum in

[21]It turns out that emissions during the period 2016-2024 are lower (higher) in the tax simulation than in the no-tax simulation when ϕ is above (below) 0.06.

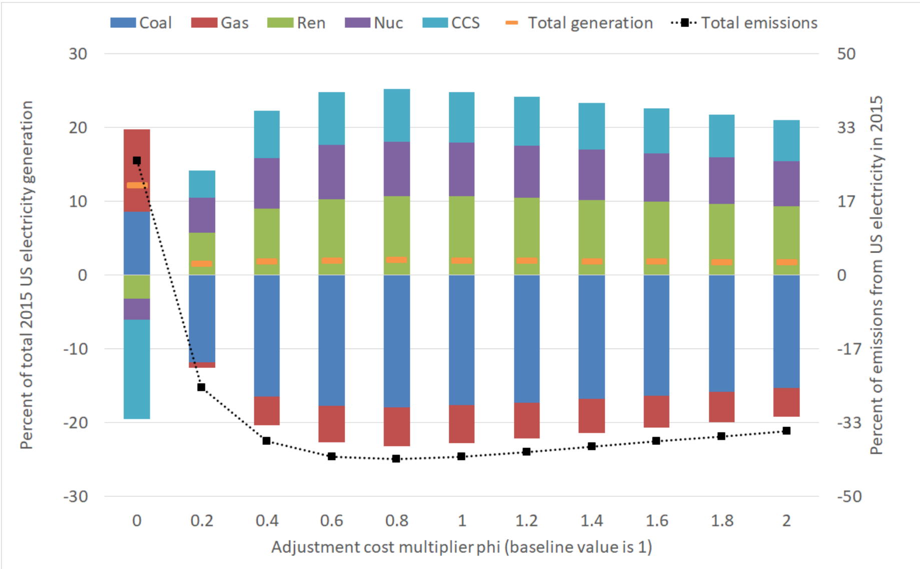

Figure 3: Sensitivity wrt. adjustment costs. Effects of tax announcement on production by source (left axis) and emissions (right axis) summed over the 10 years before the tax is implemented. Tax minus no tax simulation values.

2050 was reduced to a mere 84 TWh in the simulation without emission taxes and five-fold resource scarcity parameter values. Note that higher early gas production, caused by announcement of future emission taxes, may crowd out electricity supply from emission intensive coal plants along with electricity from low emission sources.

Last, the results are very robust to changes in initial production capacity between energy sources. Specifically, emissions during the period 2016-2024 declines following tax announcement even when initial capacity is adjusted such that all electricity in 2015 are generated from gas and petroleum fired power plants. 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 Transition dynamics under optimal versus Pigou taxes

In this Section I compare model dynamics under optimal taxes (cf. Proposition 2) and standard Pigouvian taxes on emissions equal to the social cost of carbon (henceforth referred to as a 'SCC tax'). I assume that $\psi = 0$

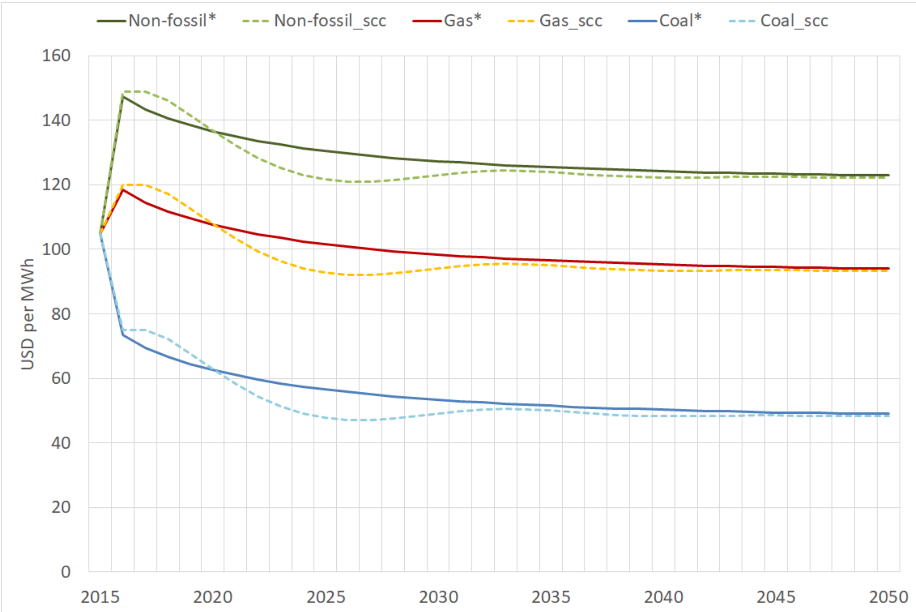

Figure 4: Selected producer prices from 2015 to 2050. Optimal (*) and SCC (_scc) trajectories.

and $\vartheta = 1/2$ in Section 3.2, implying that the firm's expectations are perfectly adaptive with quite fast error correction (see Fig 10 in Appendix B). I model CO2 emissions as a flow pollutant with constant marginal environmental damage, such that the 80 percent reduction target is reached in 2050 along the socially optimal time trajectory. Constant marginal damage is a reasonable approximation for a global pollutant like CO2.[22]

Figure 4 graphs selected producer prices over the period 2015 to 2050. The taxes are introduced in 2016, which is the first year in the simulation runs. The consumer prices are roughly one USD per MWh above the producer prices for non-fossil energy (nuclear and renewables, which are one top of each other in Figure 3) over the whole time horizon. The tax implementation induces a sharp increase in the supply cost of electricity from coal fired power plants in 2016. This also increases residual demand and, hence, producer prices for electricity from low emission power plants. The prices decrease after 2016, because increased generation capacity from non-fossil energy sources replace the fossil fuels with relatively high supply costs

[22]See Appendix B for results with a stock pollutant with a yearly depreciation rate of 0.5 percent, implying a CO2 half-life of 139 years. The qualitative results presented in Section 3.2 are not affected by the inclusion of stock dynamics.

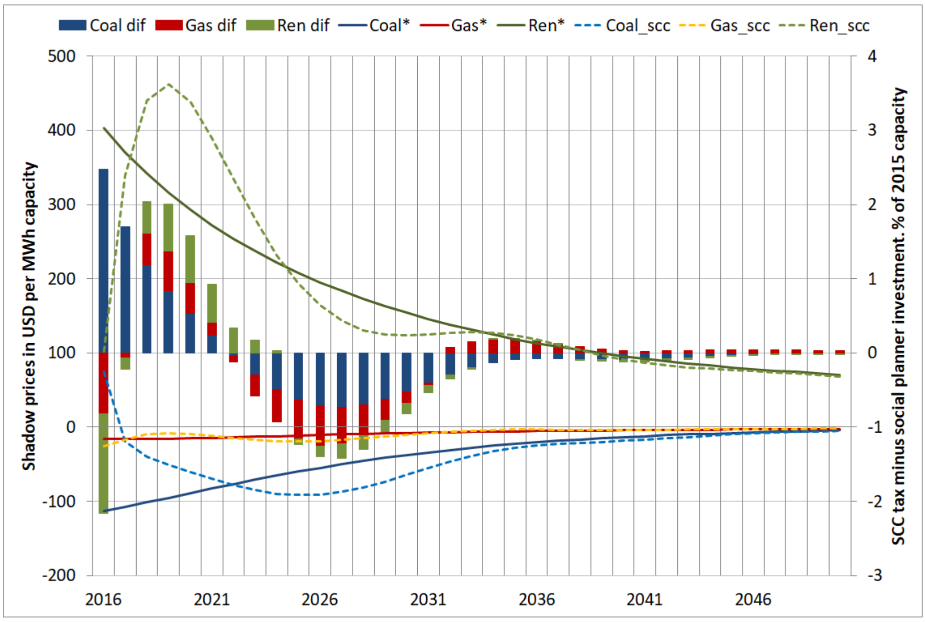

Figure 5: Selected shadow prices on capacity (line chart) and differences in net investment levels (bargraph).

(because of the emission tax).

Whereas the prices decline steadily after 2016 along the socially optimal time trajectory, prices in competitive equilibrium with adaptive expectations and SCC taxes oscillate. The reason is that the adaptive price expectations lags actual prices, which tends to cause too high shadow prices on capacity in periods with decreasing producer prices, and too low shadow prices in periods with increasing producer prices. These mechanics are revealed in Figure 5, which graphs selected shadow prices and net investment levels over the period 2016 to 2050. We observe that the low emission electricity generation sources adjust their price expectations to the new and higher price levels after around two years. Therefore, we have too little investment in low emission capacity in the first two simulation years. Further, the firms do not foresee that the equilibrium prices will decrease as the economy completes the transition towards the new low emission electricity market. Consequently, the 'short run' dynamics under SCC taxes are characterized by overinvestment in coal fired power plants, and an expansion of low-emission generation capacity that catches up and 'overshoot' after two years of underinvestment. The resulting excess capacity from all generation

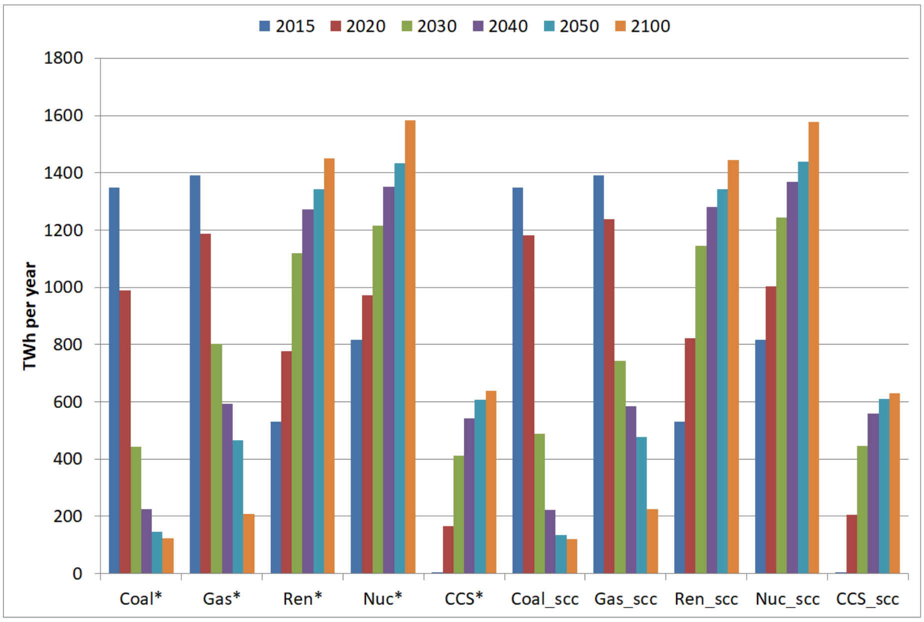

Figure 6: U.S. electricity production by source. Optimal (*) and SCC (_scc) trajectories. Selected years.

sources causes the equilibrium prices to decline, and consumer prices under the SCC tax are below those of the optimal tax in the 'medium run' (i.e., rougly the 2020s and 2030s).[23] In the 'long run', the time trajectory under SCC taxes converge towards the socially optimal path as the adaptive expectations converge towards the actual prices realized in competitive equilibrium.

Figure 6 graphs electricity production levels by energy source under optimal taxes and SCC taxes in selected years, including historic EIA numbers for the year 2015. The total present value welfare gain following implementation of optimal versus SCC taxes constitutes 32 percent of total supply costs in 2015. The welfare effects are graphed in Figure 13 in Appendix B.

Whereas the presence of adaptive expectations implies that the transition towards the low carbon economy under Pigouvian taxes is too slow in the first years, the results are less clear in the slightly longer run (both in theory and in the numerical model). The reason is that the adaptive expectations adjusts slowly to the decreasing prices after the tax is implemented (see

[23]Note that these mechanics are related to the well-known cobweb model, where the amount produced must be chosen before prices are observed, see e.g. Ezekiel (1938).

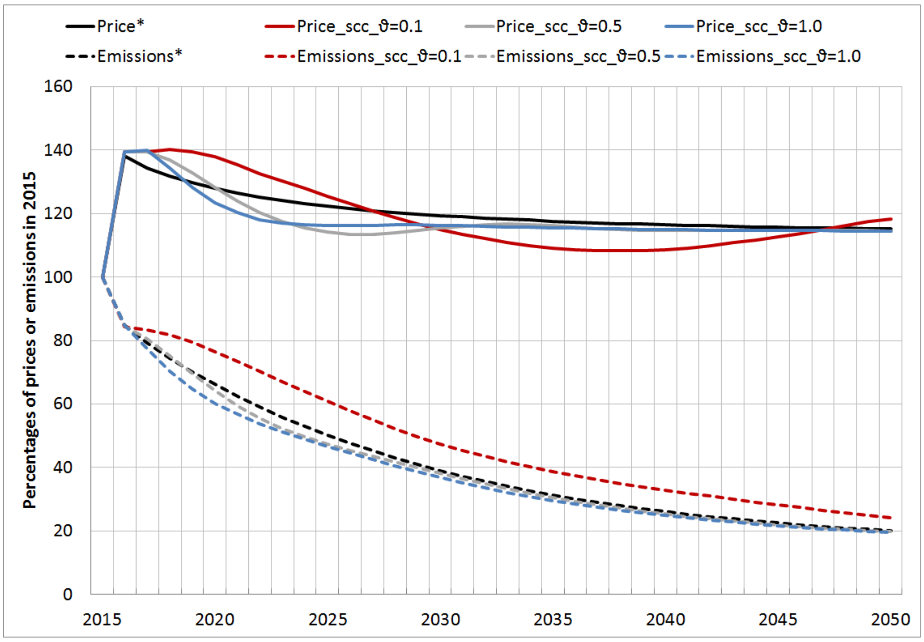

Figure 7: Consumer prices and emissions in percent of 2015 US electricity price and 2015 US emissions from electricty generation, respectively. Optimal and SCC taxes with different values on the error correction parameter ϑ.

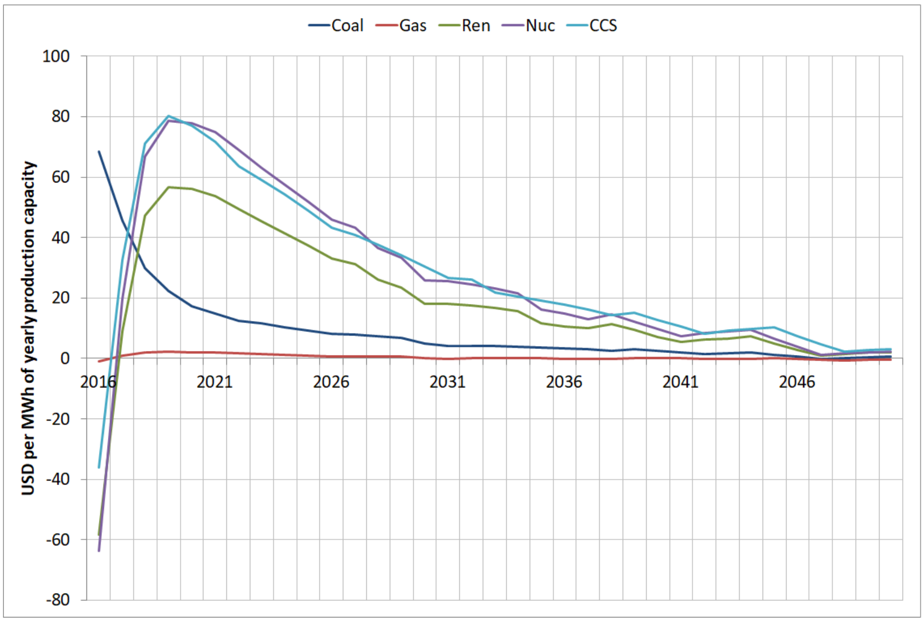

Figure 8: Optimal taxes on investment by energy source, $\varphi_t^{i,*}$.

figure 10 in Appendix B), implying a period with overinvestment in all fuel sources (see Figure 5). Hence, it is possible for emissions during, e.g., the first decade or two, to be either too high or too low under Pigouvian taxes. Figure 7 graphs consumer prices and total emissions under optimal taxes and SCC taxes, with various adaptive expectation error correction parameter values. Whereas emissions under SCC taxes are above the optimal levels with slow updating of price expectations (cf., $\vartheta = 0.1$ in Figure 7), rapid updating of beliefs (cf., $\vartheta = 1$ in Figure 7) causes investments in low emission technology to overshoot, with associated emissions below the optimal time trajectory. The baseline simulation error correction parameter value (cf., $\vartheta = 0.5$ in Figure 7) is somewhere in between. Hence, the role of the optimal taxes given in Proposition 2 depends crucially on the expectations formation process (11). Specifically, the role of the optimal taxes is to speed up the transition towards the low carbon economy if and only if expectations adjust slowly to new information.

The optimal taxes on investment in the baseline simulation, $\varphi_t^{i,*}$, are graphed in Figure 8. Note that these taxes are not simply the difference between the shadow prices on investment depicted in Figure 5, because current taxes influence the firms' expectations about future producer prices.

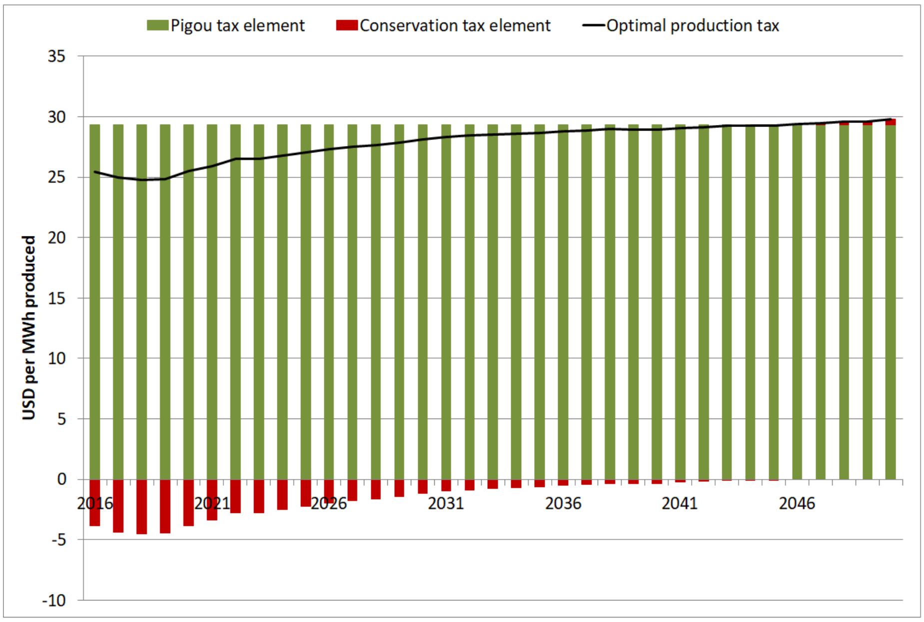

Figure 9: Optimal tax on electricity from gas fueled power plants, $\tau_t^{gas,*}$.

The optimal investment taxes are positive (or close to zero) for all fuels in all periods, except for a subsidy to low emission energy in the first simulation year 2016. Somewhat surprisingly, perhaps, the investment taxes are higher for low emission energy after 2017. The reason is that these energy sources experience a stronger discrepancy between expected and actual producer prices than the fossil fuels coal and gas.

Figure 9 graphs the optimal production tax for electricity from gas fueled power plants. We observe that the shadow price tax elements are negative over the years 2016-2045, which is consistent with declining producer prices and Proposition 2. The conservation tax elements are positive after 2045, because resource scarcity induces slowly increasing producer prices in the long run (cf., Corollary 2). The conservation tax elements are very small for the other energy sources. Remember that the optimal taxes in figures 8 and 9 are sensitive to the assumptions about the adaptive expectations formation process (11), cf. Figure 7.

4 Conclusion

This paper examined regulation in the presence of adjustment costs, resource scarcity and potentially imperfect knowledge about future prices. There are three main lessons two be learned. First, the results indicate that announcement of future environmental regulation of electricity markets will reduce current emissions. In short, whereas demand side dynamics induced by adjustment costs indicate that future taxes reduces early emissions, the supply side dynamics put forward by the green paradox literature suggests an increase in early emissions. Whereas theory alone hence yields ambiguous results, the demand side dynamics dominates the supply side dynamics for all reasonable parametrizations of the numerical model. Second, the socially optimal time path can be achieved with standard Pigouvian taxes if and only if the firms have perfect knowledge about the future. If the firms' expectations are partly adaptive, optimal taxation includes a tax on investment and, in the case of scarce resources, taxes that differs from the Pigouvian emission tax. Third, the presence of adaptive price expectations does not necessarily induce excess inertia in the transition towards a low emission energy market. Indeed, the combined presence of adjustment costs and adaptive expectations may induce overinvestment in clean production capacity and emissions below the optimal time trajectory.

The theory predicts that forward-looking agents 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.[24] 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 increasingly competitive renewable power partly explains the investors' vanishing interest in coal.[25]

Last, the paper features a stylized model framework and issues like

[24] According 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).

[25] The 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).

commitment, uncertainty, network externalities, economic growth, technical change 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.

Appendix A: Proofs and derivations

Proof of Lemma 1. The benevolent social planner solves (5) s.t. equation (1) to (4). The associated present value Hamiltonian is:

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

where $c(\cdot)$ is given by (1), and $\hat{\lambda}_t^i$, $\hat{\mu}_t^i$ and $\hat{\gamma}_t$ are shadow prices on production capacity, cumulative production (over time) and the emission stock, respectively. The Maximum principle for discrete time optimization states that the solution to (5) must satisfy the following necessary conditions for all $i \in I$ (see, e.g., Sydsæter et al., 2008, p. 445):

$$\begin{aligned} H_{x_t^i}^* &= \delta^{t-1}\left(u_{x_t^i}\left(\boldsymbol{x}_t^*\right) - c_{x_t^i}^i\left(x_t^{i,*}, X_t^{i,*}, Y_t^{i,*}\right) - d_{x_t^i}\left(\varsigma \boldsymbol{x}_t^{*\prime}, S_t^*\right)\right) + \hat{\mu}_t^i + \zeta^i\gamma_t^* \le 0, \forall t, \quad (16) \\ H_{y_t^i}^* &= -\delta^{t-1}\kappa_{y_t^i}^i\left(y_t^{i*}\right) + \hat{\lambda}_t^{i,*} = 0 \ \forall t, \\ \hat{\lambda}_{t-1}^{i,*} &= H_{Y_t^i}^* = -\delta^{t-1} f_{Y_t^i}^i\left(x_t^{i,*} - Y_t^{i,*}\right) + \beta\hat{\lambda}_t^{i,*}, \ \forall t \neq \bar{t}, \\ \hat{\mu}_{t-1}^{i,*} &= H_{X_t^i}^* = -\delta^{t-1} h_{X_t^i}^i\left(X_t^{i,*}\right) x_t^{i,*} + \hat{\mu}_t^{i,*}, \ \forall t \neq \bar{t}, \\ \hat{\gamma}_{t-1}^* &= H_{S_t}^* = -\delta^{t-1} d_{S_t}\left(\varsigma \boldsymbol{x}_t^{*\prime}, S_t^*\right) + \alpha\gamma_t^*, \ \forall t \neq \bar{t}, \\ 0 &= \hat{\lambda}_{\bar{t}}^{i,*} = \hat{\mu}_{\bar{t}}^{i,*} = \hat{\gamma}_{\bar{t}}^*, \end{aligned}$$

where the last line is the transversality conditions for free state variables $Y_{\bar{t}}^i$, $X_{\bar{t}}^i$ and $S_{\bar{t}}$. The assumptions imposed on the cost function $c(\cdot)$ ensures that the Hamiltonian is concave for all $t \in T$ around optimum. Hence, the necessary conditions (16) maximize W by Arrow's Sufficient Theorem. Last, the state movement equations (2), (3) and (4) must be satisfied along the optimal trajectory.

The solution to $\hat{\lambda}_{t-1}^{i,*} = -\delta^{t-1} f_{Y_t^i}^i\left(x_t^{i,*} - Y_t^{i,*}\right) + \beta\hat{\lambda}_t^{i,*}$ in (16) is $\hat{\lambda}_t^{i,*} =$

$\frac{\hat{\lambda}_0^{i,*}}{\beta^t} + \sum_{r=t+1}^{r=\bar{t}} \frac{\delta^{r-1}}{\beta^{t-r+1}} f^i_{Y^i_t}\left(x_t^{i,*} - Y_t^{i,*}\right)$. The transversality condition $\hat{\lambda}_{\bar{t}}^{i,*} = 0$ then implies $\hat{\lambda}_0^{i,*} = -\sum_{r=1}^{r=\bar{t}} (\beta\delta)^{r-1} f^i_{Y^i_t}\left(x_t^{i,*} - Y_t^{i,*}\right)$. Inserting in the equation for $\hat{\lambda}_t^{i,*}$ above yields $\hat{\lambda}_t^{i,*} = -\sum_{r=t+1}^{r=\bar{t}} \delta^{r-1}\beta^{r-t-1} f^i_{Y^i_t}\left(x_t^{i,*} - Y_t^{i,*}\right)$ $(t < \bar{t})$. The current value shadow price on capacity is then given by:

$$\lambda_t^{i,*} \equiv \frac{\hat{\lambda}_t^{i,*}}{\delta^{t-1}} = -\delta \sum_{r=t+1}^{r=\bar{t}} (\beta\delta)^{r-t-1} f^i_{Y^i_r}\left(x_t^{i,*} - Y_t^{i,*}\right), \tag{17}$$

with $\gamma_{\bar{t}} = 0$.

The solution to $\hat{\mu}_{t-1}^{i,*} = -\delta^{t-1} h^i_{X^i_t}\left(X_t^{i,*}\right) x_t^{i,*} + \hat{\mu}_t^{i,*}$ in (16) is $\hat{\mu}_t^{i,*} = \hat{\mu}_0^{i,*} + \sum_{r=1}^{r=\bar{t}} \delta^{r-1} h^i_{X^i_r}\left(X_r^{i,*}\right) x_r^{i,*}$ for $t < \bar{t}$. The transversality condition $\hat{\mu}_{\bar{t}}^{i,*} = 0$ then implies $\hat{\mu}_0^{i,*} = -\sum_{r=1}^{r=\bar{t}} \delta^{r-1} h'_{X^i_r}\left(X_r^{i,*}\right) x_r^{i,*}$. Hence, we have $\hat{\mu}_t^{i,*} = -\sum_{r=t+1}^{\bar{t}} \delta^{r-1} h'_{X^i_r}\left(X_r^{i,*}\right) x_r^{i,*}$ $(t < \bar{t})$. The current value shadow price on cumulative production $X_t^{i,*}$ is then given by:

$$\mu_t^{i,*} \equiv \frac{\hat{\mu}_t^{i,*}}{\delta^{t-1}} = -\sum_{r=t+1}^{\bar{t}} \delta^{r-t} h'_{X^i_r}\left(X_r^{i,*}\right) x_r^{i,*},\ t < \bar{t}, \tag{18}$$

with $\mu^i_{\bar{t}} = 0$.

The solution to $\hat{\gamma}^*_{t-1} = H'_{S_t} = -\delta^{t-1} d_{S_t}(\boldsymbol{\varsigma x}_t^{*\prime}, S_t^*) + \alpha\hat{\boldsymbol{\gamma}}_t$ in (16) is $\hat{\gamma}^*_t = \frac{1}{\alpha^t}\hat{\gamma}_0 + \sum_{r=1}^{r=t} \frac{\delta^{r-1}}{\alpha^{t-r+1}} d_{S_r}(\varsigma x_r, S_r)$. The transversality condition $\hat{\gamma}_{\bar{t}} = 0$ then implies $\hat{\gamma}_0 = -\alpha^{\bar{t}} \sum_{r=1}^{r=\bar{t}} \frac{\delta^{r-1}}{\alpha^{t-r+1}} d_{S_r}(\varsigma x_r, S_r)$. Inserting in the equation for $\hat{\gamma}_t$ above yields $\hat{\gamma}_t = -\sum_{r=t+1}^{\bar{t}} \delta^{r-1}\alpha^{r-t-1} d_{S_r}(\varsigma x_r, S_r)$ $(t < \bar{t})$. This is the adjoint related to the emissions stock S_t in the maximization problem (5). The solution in Lemma 1 is presented in terms of the marginal environmental damage from current emissions (or social cost of carbon), however, which is then given by:

$$\gamma_t \equiv d_{(\boldsymbol{\varsigma x}_t^{z\prime})}\left(\boldsymbol{\varsigma x}_t^{z\prime}, S_t^z\right) + \frac{-\hat{\gamma}_t}{\delta^{t-1}} = d_{(\boldsymbol{\varsigma x}_t^{z\prime})}\left(\boldsymbol{\varsigma x}_t^{z\prime}, S_t^z\right) + \delta \sum_{r=t+1}^{r=\bar{t}} (\alpha\delta)^{r-t-1} d_S(\boldsymbol{\varsigma x}_t^{z\prime}, S_r)_r,\ t < \overline{t}, \tag{19}$$

with $\gamma_{\bar{t}} = d_{\left(\boldsymbol{\varsigma x}_{\bar{t}}^{z\prime}\right)}\left(\boldsymbol{\varsigma x}_{\bar{t}}^{z\prime}, S_{\bar{t}}^z\right)$.

Inserting equations (17) to (19) in the first line in equation (16) yields the socially optimal time trajectory in Lemma 1.

Proof of Lemma 2. The representative firm maximizes the present value of profits over the whole time horizon, given its beliefs about future prices. The period t Hamiltonian of firm i is:

$$H_s^{i,t} = \begin{cases} \delta^{s-1}\left[\left(p_s^{i,t,e} - \tau_s^{i,t,e}\right) x_s^{i,t} - c^i\left(x_s^{i,t}, X_s^{i,t}, Y_s^{i,t}\right) - \left(\kappa^i\left(y_s^{i,t}\right) + \varphi_s^{i,t,e} y_s^{i,t}\right)\right] \\ \quad + \hat{\lambda}_s^i\left(\beta Y_s^{i,t} + y_s^{i,t}\right) + \hat{\mu}_s^i\left(X_s^{i,t} + x_s^{i,t}\right), \ \forall t < \bar{t}, \\ \delta^{\bar{t}-1}\left[\left(p_{\bar{t}}^{i,t,e} - \tau_{\bar{t}}^{i,t,e}\right) x_{\bar{t}}^{i,t} - c^i\left(x_{\bar{t}}^{i,t}, X_{\bar{t}}^{i,t}, Y_{\bar{t}}^{i,t}\right) - \left(\kappa^i\left(y_{\bar{t}}^{i,t}\right) + \varphi_{\bar{t}}^{i,t,e} y_{\bar{t}}^{i,t}\right)\right], \end{cases}.$$

for all $s \in \xi_t = \left\{t, t+1, ..., \bar{t}\right\}$. Note that the variables for the time t trajectory have superscript t, because the trajectory is updated in the next period unless $\psi = 0$ (or if the economy is in a stationary state and $\vartheta = 1$). The system of necessary conditions is (the derivation is similar to the derivation of the social planner's time trajectory above):

$$\begin{aligned} p_s^{i,e,t} - \tau_s^{i,e,t} &\leq c_{x_s^i}^i\left(x_s^{i,t}, X_s^{i,t}, Y_s^{i,t}\right) - \mu_s^{i,t}, \\ \lambda_s^{i,t} &\leq \kappa_{y_s^i}\left(y_s^{i,t}\right) + \varphi_s^{i,t}, \\ \lambda_s^{i,t} &= -\delta \sum_{r=s+1}^{r=\bar{t}} (\beta\delta)^{r-s-1} f_{Y_t^i}^i\left(x_r^{i,t} - Y_r^{i,t}\right), \\ \mu_s^{i,t} &= -\sum_{r=s+1}^{\bar{t}} \delta^{r-s} h_{X_r^i}^i\left(X_r^{i,t}\right) x_r^{i,t}, \end{aligned} \tag{20}$$

which together with the state movement equations (2) and (3) is sufficient for optimum by Arrow's Sufficient Theorem (given the firm's beliefs about future prices and taxes). The sums over r are zero at $t = \bar{t}$. Note that the control sequence $\left\{x_s^{i,t}, y_s^{i,t}\right\}_{s\in\xi_t}$ is a function of the prices p_t^i, $\tau_t^i$and φ_t^i (including expectations about the future), and the state variables X_t^i and Y_t^i. This allows the firm to update its trajectory based on the latest available information about the current state of the system (i.e., $\left\{x_s^{i,t}, y_s^{i,t}\right\}_{s\in\xi_t}$ is a closed-loop or Markov control). Market equilibrium requires that $p_t^i = u_{x_t^i}\left(x_t^i\right)$ in (20) along the competitive equilibrium path (cf. equation 9).

The rational expectations time t trajectory $\left\{x_s^{i,t,rat}, y_s^{i,t,rat}\right\}_{s\in\xi_t}$ solves the maximization problem (10) subject to equations (2), (3) and (9) under the assumption of perfect information about the future along the rational expectations path (which coincides with the competitive equilibrium iff $\psi^1 =$

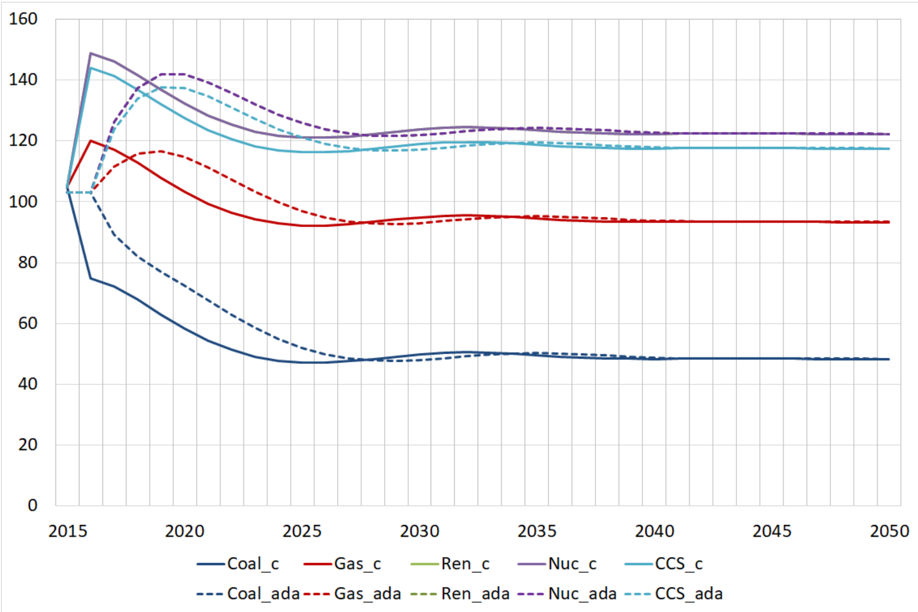

Figure 10: Producer prices in competitive equilibrium. Actual (c) and one period ahead adaptive expectations (ada).

1) for all $i \in I$. The solution is given by (20), with perfect information about future prices along the rational expectations path.

The adaptive expectations time t trajectory $\left\{x_s^{i,t,ada}, y_s^{i,t,ada}\right\}_{s\in\xi_t}$ solves the maximization problem (10) subject to equations (2), (3) and (11). The solution to (11) is $\chi_{t+n}^{i,e,t} = (1-\vartheta)^t \bar{\chi}^i + \vartheta \sum_{k=1}^t (1-\vartheta)^{t-k} \chi_k^{i,c}$. Hence, the adaptive expectations trajectory is given by (20) with $p_s^{i,e,t} - \tau_s^{i,e,t} = (1-\vartheta)^t \left(\bar{p}^i - \bar{\tau}^i\right) + \vartheta \sum_{k=1}^t (1-\vartheta)^{t-k} \left(p_k^i - \tau_k^i\right)$ and $\varphi_s^{i,e,t} = (1-\vartheta)^t \bar{\varphi}^i + \vartheta \sum_{k=1}^t (1-\vartheta)^{t-k} \bar{\varphi}_k^i$ for $s > t$. We have $p_t^{i,e,t} = p_t^i$, $\tau_s^{i,e,t} = \tau_t^i$ and $\varphi_t^{i,e,t} = \varphi_t^i$ at $s = t$ in (20); cf., equations (12) and (13) in the text.

We have to solve for the rational expectations trajectory and the adaptive expectations trajectory in each period $t < \bar{t}$ in order to find the current competitive equilibrium production and investment levels, unless $\psi = 1$, because the representative firm updates its beliefs about the future based on the latest available information. Lemma 2 now follows from equations (9) and (15).

Appendix B: The numerical model

Let the set of goods be $I = \{coal, gas, nuclear, renewables, CCS\}$, such that x_t^i denotes U.S. electricity produced (and consumed) in year $t \in T$ from energy source $i, j \in I$ (j is alias). The utility function is given by $u(x_t) = u_1\left(\sum_{i \in I} x_t^i\right) - (u_2/2)\left(\sum_{i \in I} x_t^i\right)^2$. The 'standard' part of the cost function is $k^i(x_t^i) = c_1 x_t^i + \left(c_2^i/2\right)\left(x_t^i\right)^2$, where c_1 and c_2^i are fuel specific calibrated parameters. The adjustment cost function is $f^i\left(\left(x_t^i - Y_t^i\right)/\frac{1}{2}\left(\sum_i Y_t^i/5 + Y_t^i\right)\right) = g^i(\cdot)\left(c_4^i + \left(1 - c_4^i\right)C^i(\cdot)\right)$,with $g^i(\cdot) = (c_3/2)\left(\left(x_t^i - Y_t^i\right)/\frac{1}{2}\left(\sum_i Y_t^i/5 + Y_t^i\right)\right)^2$ and $C^i(\cdot) = \frac{1}{\pi}\sum_{i \in I}\left[arctan\left(\left(x_t^i - c_5 x_{t-1}^i\right)/c_6\right) + \frac{1}{2}\right]$. Here c_3 determines the magnitude of the adjustment costs, c_4^i is the share of adjustment costs that is incurred when production is declining, c_5 is the capital depreciation factor, and c_6 determines the shape of $C^i(\cdot)$. The function $C^i(\cdot)$ is derived using the Cauchy cumulative distribution function. Note that $C^i(\cdot) \in (0,1)$ and increases steeply from near zero to near 1 around $x_t^i - c_5 x_{t-1}^i = 0$, given a low value on c_6. Figure 11 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 function $h^i\left(X_t^i\right) = c_7^i X_t^i$.[26] 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, and calibrated under the assumption that half of the coal is available for CCS).

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.[27] I let the

[26] NEB: 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

[27] I 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:

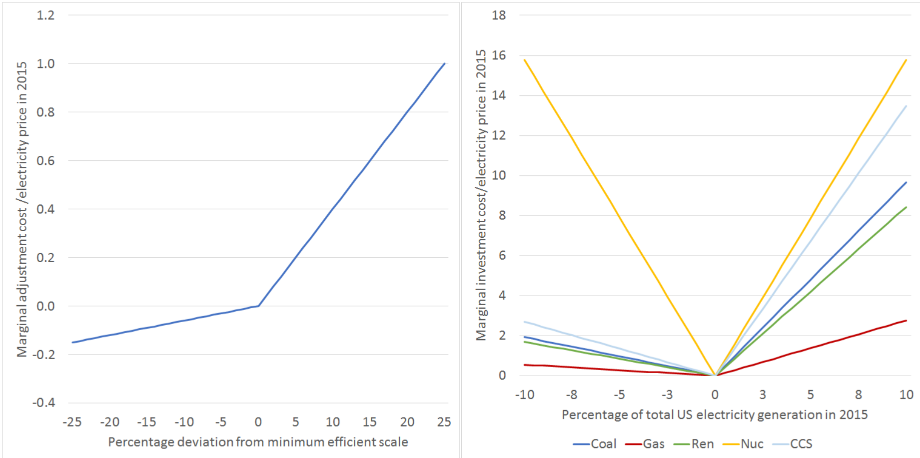

Figure 11: Calibrated adjustment and investment costs. $f_x^i\left(x_t^i - Y_t^i\right)$ on the left and $k_x^i\left(x_t^i\right)$ on the right.

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 two equation system is: (i) $Elcons = 1806 - 2.99208 * Elprice + 131.0206 * GDP + 13.50116 * Gasprice$, and (ii) $Elprice = 197.0943 - 0.0325257 * Elcons + 0.313119 * Oilprice$. Here electricity consumption (Elcons) is measured in TWh, GDP is in trillions of USD (2014), electricity prices (Elprice) are in USD (2014) per MWh, gas prices are in USD (2014) per million Btu, and oil prices are USD (2014) per barrel. All variables are significant at a 5 percent confidence level and the R^2 values are 0.9864 and 0.8790 for equations (i) and (ii), respectively. Note the negative sign on electricity consumption (Elcons) in equation (ii). Alternative estimations featuring the real interest rate, wage index and U.S. coal prices give very similar results. One lag Dickey-Fuller unit root test suggests 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 (i). Environmental damage is given by $d\left(\varsigma \boldsymbol{x}_t', S_t\right) = d_1 \varsigma \boldsymbol{x}_t' + d_2 S_t + \frac{d_3}{2} S_t^2$.

Federal reserve (https://www.federalreserve.gov/releases/h15/data.htm); Inflation: (http://www.usinflationcalculator.com/inflation/historical-inflation-rates/).

The derivation of the other parameters are described in Section 3. See Table 1 for exact parameter values.

The numerical model solves Lemmas 1 and 2 (extended to include capital deprecation), given these functional forms and parameter values. The competitive equilibrium is solved as a 'recursive loop' over the years 2015, 2016, $\ldots$,2215. In this loop, the maximization problem (10) is solved for $\bar{t} - t = 200$ in each year (i.e., $\bar{t} = 2215$ when $s = 2016$,$\bar{t} = 2216$ when $s = 2017$,..., $\bar{t} = 2415$ when $s = 2215$). The initial conditions $\bar{Y}^i$, X^i, $\bar{S}$ and $\bar{\chi}^i$, in each loop year is determined by the previous year simulation.

Figure 10 graphs competitive equilibrium producer prices and the one period ahead adaptive producer prices in the simulation run with SCC taxes. Figure 12 graphs results from an sensitivity analysis wrt. resource scarcity. Here the scarcity parameter c_7^i is multiplied with $0, 1, ..., 5$ for coal, gas and CCS. The left hand side of Figure 13 graphs the model simulated welfare effects following implementation of optimal taxes, compared to Pigou taxes. I test the model fit by running the model against history from 1991 to 2015. The right hand side of Figure 13 graphs model projections and historic figures for the period 1991-2015. This simulation uses $\psi = 1$ and features real figures for U.S. GDP and U.S. prices on coal and gas (captured by the producer tax). The simulation is very rough and does not capture other variables that may affect U.S. electricity generation. The simulation run assumes that coal and gas prices remain at the 2015 level into the future, and that future U.S. GDP grows at a rate equal to the average growth rate during 1990-2015. We see that the simulation is unable to capture dynamics between gas and coal induced by the shale gas revolution. Figure 14 replicates Figures 4 and 5 in the text in the case of a stock pollutant. This simulation uses $\alpha = 0.995$ (according to Hoel and Karp (2002), 0.5% is widely accepted as an approximate point estimate for the decay rate for greenhouse gasses), $d_1 = 1$ and $d_2 = 0.0545/1000$.

Parameter\fuel	Coal	Gas	Nuclear	renewable	CCS
ς^i	1.011	.3986	.0052	.0052	.0705
x_0^i	1349	1388	531.4	817.5	1
c_1^i	42.98	40.62	68.3	78.8	94.3
c_2^i	.0381	.0387	.0319	.0294	.0226
c_4^i	0.1	0.1	1	0.1	0.1
c_6^i	105	105	0	0	105
c_7^i	7.4E-5	9.6E-4	0	0	1.4E-4
k_1^i	.963	.276	1.574	.841	1.346
k_2^i	0.2	0.2	0.2	1	0.2
$\bar{Y}^i$	1348.8	1387.7	817.5	531.4	5
Parameter	α	δ	ϑ	ψ	c_3
Value	.995	0.97	0.5	0 or 1	4.2E+5
Parameter	c_5	d_1	d_2	d_3	k_3
Value	.994	73.4	0	0	1.0E-4
Parameter	$\bar{p}^i$	$\bar{X}^i$	u_1	u_2	
Value	103.2	1	705.5	.3342	

Table 1: Parameter values in the numerical illustration.

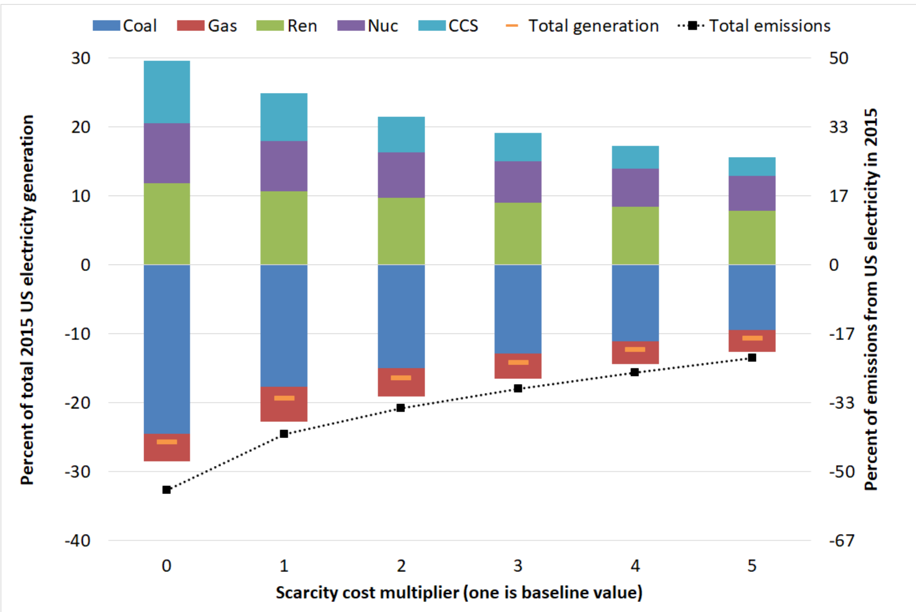

Figure 12: Sensitivity wrt. resource scarcity. Effects of tax announcement on production by source (left axis) and emissions (right axis) summed over the 10 years before the tax is implemented. Tax minus no tax simulation values.

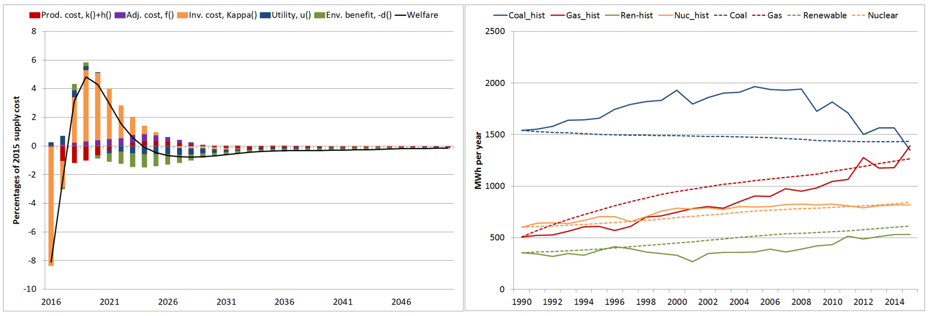

Figure 13: Left figure: Welfare effects. Optimal tax simulation values minus Pigou tax simulation values. Right figure: Model fit to history. Electricity generation by source. Model generated (dashed lines) and historic values (unbroken lines).

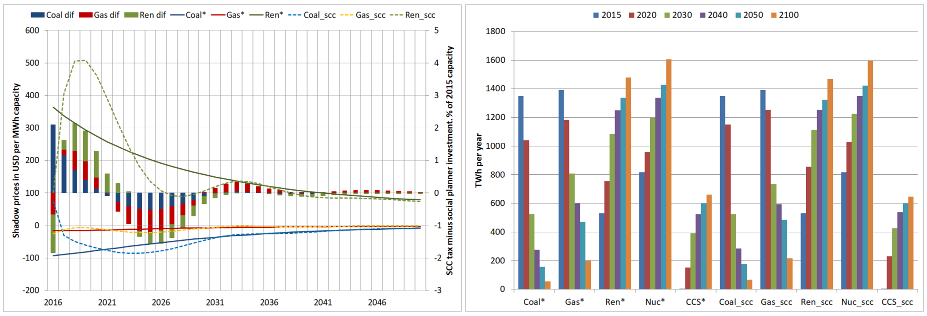

Figure 14: Transition dynamics with a stock pollutant.

References

Abel, A. B. & Eberly, J. C. (1996), 'Optimal investment with costly reversibility', Review of Economic Studies 63, 581.

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.

Chow, G. C. (1989), 'Rational versus adaptive expectations in present value models', he Review of Economics and Statistics 71(3), 376-384.

Chow, G. C. (2011), 'Usefulness of Adaptive and Rational Expectations in Economics', CEPS Working Paper No. 221, www.princeton.edu/~ceps/workingpapers/221chow.pdf

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.

Ezekiel, M. (1938), 'The cobweb theorem', The quarterly journal of economics 52(2), 255–280.

Friedman, M. (1957), Theory of the consumption function, New Jersey, USA: Princeton University Press.

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.

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. & Karp, L. (2002), 'Taxes versus quotas for a stock pollutant', Resource and Energy Economics 24(4), 367-384.

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.

IEA (2016), 'World Energy Outlook', Technical report, OECD/IEA, Paris.

IPCC (2015), 'Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.))', Technical report, IPCC, Geneva, Switzerland.

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.

Lucas, R. (1976), 'Optimal investment policy and the flexible accelerator', International Economic Review 8(1), 78-85.

Lucas, R. E. J. (1987), Models of Business Cycles, Oxford, UK: Basil Blackwell.

Muth, J. F. (1961), 'Rational expectations and the theory of price movements', Econometrica 29(3), 315-335.

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.

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.

Sargent, T. J. (1999), The Conquest of American Inflation, New Jersey, USA: Princeton University Press.

Shapiro, M. D. (1986), 'The dynamic demand for capital and labor', The

Quarterly Journal of Economics 101(3), 513-542.

Sheffrin, S. M. (1996), Rational Expectations, Cambridge UK: Cambridge University Press.

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.

Sydsæeter, K.; Hammond, P.; Seierstad, A. & Strøm, A. (2008), Further mathematics for economic analysis, Harlow, UK: Pearson Education Limited.

Statistics Norway

Postal address:
PO Box 8131 Dept
NO-0033 Oslo

Office address:
Akersveien 26, Oslo
Oterveien 23, Kongsvinger

E-mail: ssb@ssb.no
Internet: www.ssb.no
Telephone: + 47 62 88 50 00

ISSN: 1892-753X

Statistisk sentralbyrå
Statistics Norway

Design: Siri Boquist