Noname manuscript No.
(will be inserted by the editor)

Growth, pollution, policy!

Rob Hart

June 2018

Abstract As the scale of an economy increases, pollution flows tend to increase. We show that this scale effect is reversed in the long run — given unbounded productivity growth — if at least one clean technology exists and utility is a CES function of consumption and environmental quality; that is, long-run pollution flows approach zero. We clarify the intuition using a specified model in which pollution arises as a by-product of the use of natural-resource inputs, and calibrate a model which accounts for the timing of adoption of flue-gas desulfurization technology across countries. Policies boosting growth are not the enemy of long-run sustainability, but environmental regulations should not be sacrificed for the sake of growth; on the contrary they should be tightened in anticipation of future demand for environmental quality.

Keywords economic growth · pollution · natural resources · environmental Kuznets curve · free disposability

PACS O11 · O44

Acknowledgements Thanks to Frederic Ang, Arik Levinson, and Sjak Smulders for helpful comments, to seminar participants in Athens, Ascona, Umeå, and Uppsala, and to anonymous reviewers for feedback on previous versions.

1 Introduction

Over the last century we have emitted pollution to the atmosphere which has caused brain damage in our children on a staggering scale (lead), partially destroyed the upper atmosphere's ability to filter out damaging ultra-violet radiation (CFCs), acidified soils and waters over vast areas thereby severely damaging forest and aquatic ecosystems (SO_2 and NO_x), and significantly altered the global climate (CO_2, CH_4, etc.).[1] That pollution should expand

Department of Economics, Swedish University of Agricultural Sciences, Box 7013, 750 07 Uppsala, Sweden.
Tel.: +4618671734
E-mail: rob.hart@slu.se
ORCID: 0000-0003-2946-2164.

[1] For more on these pollutants and their regulation see von Storch et al. (2003), Sunstein (2007), Ellerman et al. (2000), and Stern (2008) respectively.

when economies grow is easily intelligible as the result of a scale effect. But it is less obvious whether or not there exists an equally general mechanism leading to a later, long-run, decline in pollution flows, or whether observed declines — such as in lead, CFCs and SO_2 — are isolated or temporary blips on an upward long-run global trend. In this paper we show that in an economy in which the value of production grows without bound then if clean alternatives exist they will be chosen in the long run, hence reversing the scale effect of growth on pollution and leading pollution flows to approach zero.

The paper adds to the literature on growth and the environment, and particularly the environmental Kuznets curve or EKC, in which the seminal work is by Grossman and Krueger (1991, 1995).[2] Grossman and Krueger (1991) first put forward the basic idea behind our analysis, that a scale effect causes an increase in pollution whereas higher WTP for a clean environment causes a subsequent decline. However, their focus is on empirical observations rather than theoretical model, and they show that at country level there is often a tendency for flows of individual pollutants to grow initially and then decline as GDP grows over time. In the subsequent literature the focus has remained on econometric analysis of empirical observations. Selden et al. (1999) and many others confirm the patterns found by Grossman and Krueger, but in the absence of a convincing theoretical explanation for why the pattern should be observed, its generality remains in doubt: flows of many pollutants are still increasing in many countries, and where a pollutant is decreasing, it could still turn up again. Furthermore, if we compare paths for the same pollutant across different countries, it is hard to find clear patterns: the turning point is neither at a given time, nor at a given level of per-capita GDP (see for instance Stern 2004).

Our explanation is both general — it builds on very mild assumptions about utility and production functions — and fully consistent with Stern's observations. It generalizes the explanation of Stokey (1998), whereas it is fundamentally different from those of Andreoni and Levinson (2001), Brock and Taylor (2010), Smulders et al. (2011), and Figueroa and Pastén (2015). There are two keys to our analysis, both of which are unique in the EKC literature: we derive restrictions on the properties of the production possibility frontier (PPF) over final-good production and pollution from the fact that pollution is a by-product of final-good production, and we derive restrictions on the properties of indifference curves over consumption and polluting emissions based on the assumption of a CES utility function over consumption and environmental quality and a general damage function.

By-production implies that the PPF is hump-shaped; for given technology, there is some rate of polluting emissions at which final-good production is maximized, and if (for some reason) more pollution is to be produced then the effort of doing so will actually detract from final-good production rather than boosting it further. Hence when production is low (because of low labour productivity), pollution is also low, even in the absence of environmental regulation, and environmental quality — defined as freedom from human-generated pollution in the public sphere — is high.[3] As labour productivity increases, by-production of pollution also increases, and environmental quality deteriorates.

The CES utility function implies that if consumption rises without bound while environmental quality is constant then WTP for better environmental quality rises without bound. Hence in a first-best regulated economy the price of emitting pollution increases with income growth, and firms shift round the PPF, reducing the ratio of pollution to production. Initially pollution increases nonetheless, but if there is a clean (zero-emissions) technology

[2] Panayotou (1993) coined the phrase.

[3] As Smulders (2006) put it (p.12), 'Prehistoric man could hunt many deer, but lacked the capacity to destroy the ozone layer.'

then it will gradually be adopted and pollution will approach zero in the long run. Intuitively, when we are very poor we enjoy high environmental quality despite low willingness to pay for it, because the small scale of the economy ensures that emissions are low. As income per capita expands—driven by technological progress—baseline polluting emissions expand, but abatement efforts also increase; the former effect dominates initially (when environmental quality is high but income low), but the latter takes over in the long run, such that both environmental quality and income increase as productivity increases.

The consistency of our explanation with the observations of Stern (2004) and others follows because the shape of the PPF of pollution and production varies between countries, even those on the same income level, as does the shape of the indifference curves. For instance, consider SO_2 emissions which are a by-product of the burning of coal for electricity production. Regarding the PPF, a country with cheap high-sulfur coal has higher abatement costs for SO_2 than a country with cheap natural gas; regarding the indifference curves, a country with high population density has higher WTP per capita to reduce SO_2 emissions than a country with low population density, because a given rate of emissions per capita leads to a higher atmospheric concentration in the former. In our model we rule out biased technological change and imperfect information, for clarity. If we relax this restriction then PPFs and indifference curves may also shift over time; for instance, development of fracking technology would reduce abatement costs in the coal-rich country.

The paper proceeds as follows. In Section 2 we develop the theoretical model. In Section 3 we develop a specified model which is at a similar level of generality to Stokey's model, but our model has a straightforward intuitive interpretation grounded in empirical cases; furthermore, it is significantly richer. And in Section 4 we further specify the model, showing how it can be calibrated to explain the timing of adoption of flue-gas desulfurization in six countries over a period of 46 years. In Section 5 we discuss the existing literature in depth. Section 6 concludes.

2 A general model

In this section we aim to set up a simple model in which the representation of production in the economy is consistent with pollution being a by-product of production of final goods, and in which households value environmental quality and the consumption of final goods in a straightforward way. We deliberately make technological progress neutral and preferences homothetic, ruling out the idea that technological progress and concomitant income growth might change the shape of the production possibility frontier over aggregate consumption and pollution or the indifference curves over aggregate consumption and environmental quality, changes which might give rise to changing patterns of polluting emissions. We do this not because such changes do not occur in real economies—they do—but because they may go in different directions, either favouring increases over time in pollution relative to aggregate production, or decreases. Our aim is to investigate the 'neutral' case.

2.1 Production

Consider an economy in which a representative firm makes widgets $X(t)$ and pollutants $P(t)$ using inputs of an effective labour–capital aggregate $A(t)$, which grows exogenously. Both X and P are non-negative and A is strictly positive; in the absence of economic activity, $P = 0$. Furthermore: (i) for given A both X and P are bounded above; (ii) for given A and P

there is some maximal production of X, and (iii) there are constant returns so that convex combinations of technologies can be used, and the production possibility set in (P,X) space (given A) is convex. We can therefore define

$$X = G(A,P), \tag{1}$$

where G is a function which returns the maximal value of X for given A and P; we call such values of X *weakly optimal*, and say that they are on the production possibility frontier (PPF). Since there are constant returns in A and P, for any point (P,X) which is weakly optimal when $A = A(t)$, the point (sP,sX) is weakly optimal when $A = sA(t)$, for all s. Now define $p = P/A$, $x = X/A$, and $G(1,P/A) = g(p)$. Then we have

$$x = g(p)$$

where g is a concave function, and the PPF plotted in (p,x) space—i.e. in intensive form—is invariant to changes in A.

We have thus defined a very simple economy in which growth (given by increases in productivity, labour, and capital) is neutral (or unbiased) in the sense that the relative costs of producing X and P do not change as A increases. We have thus ruled out the idea that changes in X/P could be driven by changes in the underlying technology.

Now return to widget production. To make widgets requires purposeful effort, and if no such effort is made then $X = 0$ (widgets will never be made by accident). The pollutant can also be made through purposeful effort, but in addition it may be made by accident, as a by-product of efforts to make widgets. Finally, the pollutant P is a bad, and an economic problem. These properties of the production function have direct implications for the shape of the firm's production possibility frontier over widgets and pollution, as follows.

1. Since making widgets requires purposeful effort, if all effort is devoted to making pollution p then widget production will be zero. So when p is maximized, $x = 0$, and $g(p)$ must meet the $x = 0$ axis at finite p greater than 0.
2. Since the pollutant is a problem, $g'(0) > 0$; otherwise the pollutant would not be a problem because all firms would set $p = 0$ at all times, even in laissez-faire.
3. Since p is a flow of human-made pollution, it must be possible to reduce p to zero, if nothing else by setting $x = 0$ (i.e. ceasing production altogether). Hence $g(p)$ must meet the $p = 0$ axis at $x \geq 0$. At this point we define the outputs $(0,\underline{x})$.

We have assumed that $g(p)$ is concave. If it is *strictly* concave then properties 2 and 3 imply that g has a single turning point (a maximum) at which point we define the outputs as $(\overline{p},\overline{x})$. If on the other hand $g(p)$ is flat at the maximal value of x (so there are many global maxima) then we define $(\overline{p},\overline{x})$ as the unique point such that x is a global maximum and p takes its lowest possible value consistent with maximization of x. Given these properties of the PPF, Figure 1 illustrates three possible cases. (We also define the maximum of the PPF plotted in (P,X) space (for given A) as $(\overline{P},\overline{X})$, and the point where the ppf meets the $P = 0$ axis as $(0,\underline{X})$.)

2.2 Utility

Consider now the utility function. We aim to specify a utility function which is neutral with respect to the valuation of consumption and environmental quality, without any esoteric

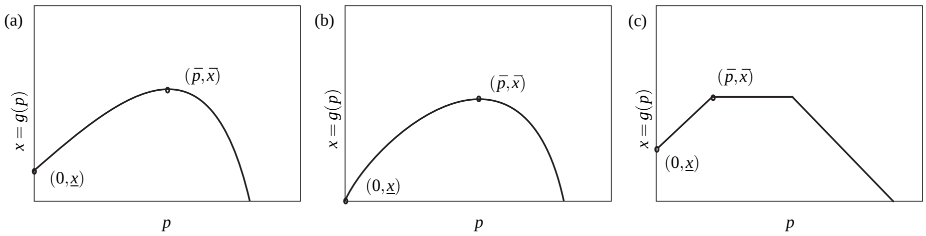

Fig. 1 The function $g(p)$ and the points $(\bar{p},\bar{x})$ and $(0,\underline{x})$, in three alternative cases: (a) strict concavity, $\underline{x} > 0$; (b) strict concavity, $\underline{x} = 0$; (c) multiple global maxima.

properties. We therefore make utility U a CES function of consumption X and environmental quality Q:

$$U = \left\{(1-\alpha)X^{(\eta-1)/\eta} + \alpha Q^{(\eta-1)/\eta}\right\}^{\eta/(\eta-1)}, \tag{2}$$

where η is the elasticity of substitution between X and Q, and $\eta > 0$. We must also link environmental quality Q to pollution P, hence we define $Q = 1/[d(P)]$, where d is differentiable and strictly increasing, and $d(0) > 0$.

As we demonstrate in Remark 1 below, this utility function encompasses both the separable (or additive) form assumed by Lopez (1994), Stokey (1998) and many others in the EKC literature, and the multiplicative form which is standard in integrated assessment models of climate policy such as Nordhaus (2008) and Golosov et al. (2014).

Remark 1 We can always write our utility function—equation (2)—in either separable or multiplicative form: separable when $\eta \neq 1$ and multiplicative when $\eta = 1$.

$$\text{Separable:} \qquad W = v(X) - h(P). \tag{3}$$

$$\text{Multiplicative:} \qquad W = X/f(P). \tag{4}$$

Furthermore, given separable preferences

$$-Xv''(X)/v'(X) = 1/\eta. \tag{5}$$

Proof There are three cases: $\eta > 1$, $\eta = 1$, and $\eta < 1$. In each case, rearrange equation 2 and $Q = 1/[d(P)]$ to obtain the results. When $\eta > 1$, we have $W = U^{(\eta-1)/\eta}/(1-\alpha)$, $v(X) = X^{(\eta-1)/\eta}$, and $h(P) = -[\alpha/(1-\alpha)]/[d(P)]^{(\eta-1)/\eta}$. When $\eta < 1$, we have $W = -U^{-(1-\eta)/\eta}/(1-\alpha)$, $v(X) = -X^{-(1-\eta)/\eta}$, and $h(P) = -[\alpha/(1-\alpha)][d(P)]^{(1-\eta)/\eta}$. And when $\eta = 1$, we have (by l'Hôpital's rule) $W = U^{1/(1-\alpha)}$, and $f(P) = [d(P)]^{\alpha/(1-\alpha)}$. Equation 5 follows directly from the expressions for $v(X)$.

Since $U'_X > 0$, we can also define the equation for the indifference curves in (X,P) space:

$$X = V(U,P) = \left[U^{(\eta-1)/\eta}/(1-\alpha) - \alpha/(1-\alpha)\cdot d(P)^{(1-\eta)/\eta}\right]^{\eta/(\eta-1)}. \tag{6}$$

Since $V'_U > 0$, higher indifference curves indicate higher utility. In addition we highlight three properties regarding V'_P, which we denote the price of pollution, since it is the optimal tax on polluting emissions in a market economy when the price of the consumption good is normalized to unity:

1. When $X \to 0$, $V'_P \to 0$ for all P, so the price of pollution is zero when consumption is zero;
2. V'_P increases monotonically in X for any $P > 0$, so the price of pollution increases with consumption;
3. When $X \to \infty$, $V'_P \to \infty$ (as long as $P > 0$), so the price of pollution approaches infinity when consumption approaches infinity.

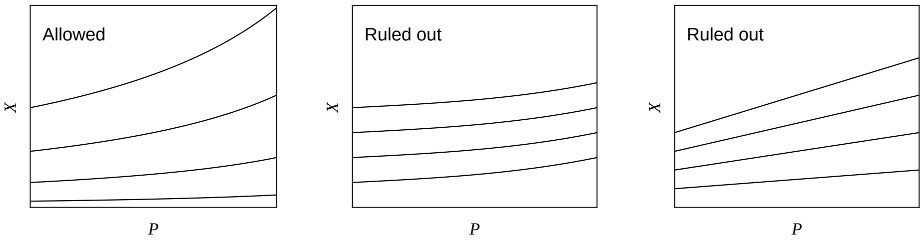

Fig. 2 Three sets of indifference curves. The second is ruled out because $\mathrm{d}X/\mathrm{d}P$ does not increase in X, implying that the WTP to remove a unit of pollution does not increase in income, and the third is ruled out because the curves are not strictly convex.

Finally, we want to ensure that $V(U,P)$ is strictly convex, which (given that $G(A,P)$ is concave) will guarantee a unique solution to the problem of maximizing utility at given A. From the expressions from the proof of Remark 1, the conditions for strict convexity are $d''/d' - d'/d[1-\alpha/(1-\alpha)] > 0$ when $\eta = 1$, and $d''/d' + (1/\eta - 2)d'/d > 0$ when $\eta \neq 1$.[4] See Figure 2 for an illustration.

2.3 The EKC

We now turn to the evolution of polluting emissions over time. Staying with our approach of building the simplest possible framework, we assume a first-best solution (achieved either through management by a social planner, or because an optimal tax on polluting emissions is imposed in an economy with otherwise perfect markets). Note also that since the PPF is concave and the indifference curves strictly convex, there is always a unique solution to the utility maximization problem for given A; the optimal choice of (P,X) is given by the point at which the highest possible indifference curve is tangent to the PPF.

Our first step is to show that over time the optimal point moves to the left around the curve $g(p)$, starting arbitrarily close to $(\overline{p},\overline{x})$, and taking a limiting value of $(0,\underline{x})$ as $t \to \infty$ (Lemma 1).

Lemma 1 *As A increases the locus of the optimal point (p,x) moves to the left along the curve $g(p)$. Furthermore,* $\lim_{A(0)\to 0}(p,x) = (\overline{p},\overline{x})$*, and* $\lim_{t\to\infty}(p,x) = (0,\underline{x})$.

Proof For the full proof see A. The essence of the proof is that the slope of the indifference curve at the point of tangency increases from 0 at $(\overline{p},\overline{x})$, approaching infinity. Hence in (p,x) space the optimal point moves to the left round $g(p)$.

[4] Note that convexity of the indifference curves is intuitively reasonable but somewhat restrictive; for instance, if the effect of some pollutant is limited to a valuable but non-essential public good then marginal damages of pollution may increase up to the point where that good is destroyed, but beyond that point marginal damages must be zero.

Based on Lemma 1 we denote $(\bar{p},\bar{x})$ and $(0,\underline{x})$ as the initial and final limits respectively. In order to investigate the implications of the movement of the optimal point from the initial to the final limit we define the elasticity of substitution between A and P—which we denote σ—and then show in Lemma 2 how σ varies as we move round $g(p)$. Finally we show in Lemma 3 how the relative sizes of σ and η determine the direction of change in P.

Definition 1 The *elasticity of substitution* between A and P at the optimal allocation (P,X) is σ, hence—since G is constant returns—(Hicks 1932) shows us that

$$\sigma = G'_P G'_A/(G''_{AP}G). \tag{7}$$

Furthermore, in intensive form we have

$$\sigma = -g'(p)\frac{g(p)-g'(p)p}{g(p)g''(p)p}. \tag{8}$$

Definition 2 A *clean technology* exists if the PPF is made up of a convex combination of n technologies (PPFs), up to $n-1$ of which are intrinsically polluting such that the PPFs meet the origin (i.e. $\underline{x}=0$), and at least one of which is intrinsically clean, such that the PPF in (p,x) space consists of a single point $(0,\underline{x})$ where $\underline{x}>0$.

Lemma 2 *The evolution of σ along $g(p)$.*

(i). *The initial limit:* $\lim_{p\to\bar{p}}\sigma = 0$.
(ii). *Moving left along $g(p)$ from $(\bar{p},\bar{x})$ σ may take any positive value; indeed, it may rise and fall between* 0 *and infinity any number of times.*
(iii). *The final limit: If a clean technology exists then there exists some non-zero level of p, which we denote $p^{\dagger}$, such that for all $p<p^{\dagger}$, $1/\sigma=0$.*

Proof See Appendix A.

Lemma 3 *In an economy as defined by equations 1 and 2: (i)* $\mathrm{d}P/\mathrm{d}A>0$ *if* $\sigma<\eta$*; (ii)* $\mathrm{d}P/\mathrm{d}A=0$ *if* $\sigma=\eta$*; and (iii)* $\mathrm{d}P/\mathrm{d}A<0$ *if* $\sigma>\eta$.

Proof See Appendix A.

The intuition behind Lemma 3 is that when consumers are inflexible (low η, a strong preference for holding X/Q constant) but firms flexible (high σ, they can easily substitute between A and P), when A rises, driving up X, inflexible consumers demand an increase in Q, and flexible firms reduce P. But when firms are inflexible and consumers flexible, when A and X rise, firms increase P and consumers accept lower Q. It is closely related to Proposition 4 of Figueroa and Pastén (2015).[5]

We now turn to the evolution of P over time: Proposition 1.

Proposition 1 *The EKC.*

(i). *There always exists some level of A, denoted $A^{\dagger}$, such that as long as $A<A^{\dagger}$, pollution P rises monotonically as A rises.*
(ii). *Assume $A(0)<A^{\dagger}$. Then an EKC is observed (i.e. there exists some time T beyond which pollution declines monotonically) if there exists some non-zero level of p, which we denote $p^{\dagger}$, such that for all $p<p^{\dagger}$, $\sigma>\eta$.*

[5] The key difference is that we have a less restrictive definition of the link between P and Q, pollution flows and environmental quality.

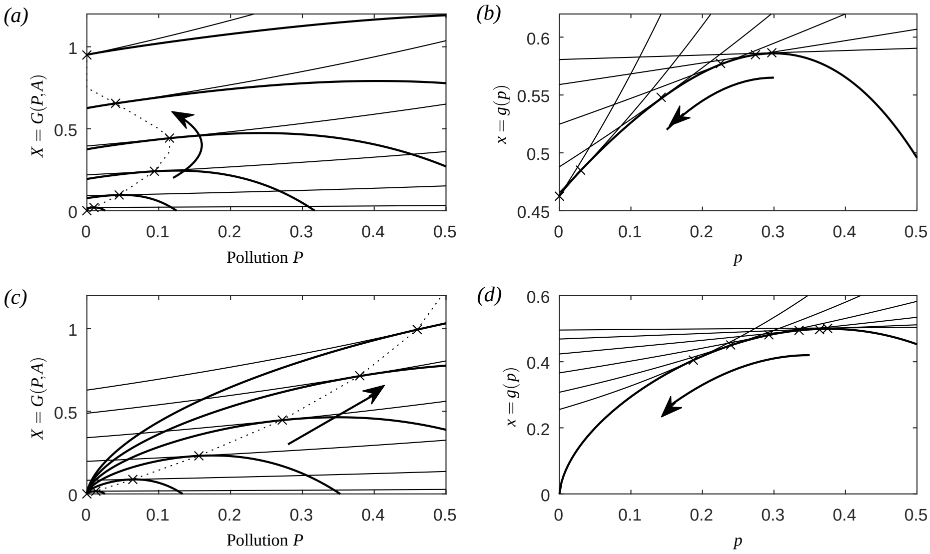

Fig. 3 The movement of (P^*,X^*) (panels (a) and (c)) and (p^*,x^*) (panels (b) and (d)) as A increases, given $\eta = 1$ and alternative ppfs. In (a) and (b) the ppf is such that $\lim_{p\to 0}\sigma > 1$, whereas in (c) and (d) $\lim_{p\to 0}\sigma < 1$. In both cases the optimal point (p^*,x^*) moves to the left in (p,x) space, in accordance with Lemma 1. But only in the upper case (when, in the limit, $1/\eta > 1/\sigma$) do we observe an EKC; in the lower case, pollution increases indefinitely. See Proposition 1.

(iii). Assume $A(0) < A^\dagger$. Then if a clean technology exists an EKC is observed for any η, and pollution $P \to 0$ as $t \to 0$.

Proof (i) Let $A(0)$, the initial quantity of labour–capital, approach zero. Then Lemma 1 tells us that $(p,x) \to (\overline{p},\overline{x})$, Lemma 2(i) tells us that $\sigma \to 0$, and hence that there must be a value of $A(0)$ below which $\sigma < \eta$, and Lemma 3 tells us that beyond this point $\mathrm{d}P/\mathrm{d}A > 0$.

(ii) Assume that there exists some non-zero level of p, which we denote $p^\dagger$, such that for all $p < p^\dagger$, $\sigma > \eta$. Then Lemma 1 tells us that as $A \to \infty$, p must fall below $p^\dagger$ in finite time, and hence that after that time (by assumption) $\sigma > \eta$. And Lemma 3 then tells us that pollution declines monotonically after that time.

(iii) If a clean technology exists then Lemma 2 shows that there must exist some $p^\dagger$ such that $1/\sigma = 0$ (and hence $\sigma > \eta$) for all $p < p^\dagger$. Furthermore, since $\mathrm{d}G/\mathrm{d}P$ is well defined at $P = 0$ the point of tangency between the indifference curve and the PPF must approach $P = 0$, hence $P \to 0$.

The intuition here is straightforward. Since the initial willingness to pay for improved environmental quality (and hence reduced pollution) is arbitrarily low, firms produce arbitrarily close to $(\overline{p},\overline{x})$, and as A increases over time, both P and X increase at the same rate. As productive potential increases and agents care little about pollution, production expands, the chosen technology stays almost the same, and pollution expands. But when income increases without bound and pollution is non-zero, WTP to reduce pollution increases without bound, and since the unit cost of switching to the clean technology is bounded above, all production will be switched to the clean technology in the limit.[6]

[6] In between the extremes, along the straight sections of the PPF marginal abatement costs do not change as abatement increases. Hence as income increases and the indifference curves steepen, the optimal point

A corollary to Proposition 1 is that when (in the limit) $\sigma = \eta$ then pollution flows approach a constant rate, and when $\sigma < \eta$ they increase.

3 A specified model

We now develop a specified model economy and derive analytical results. The main purpose is to exemplify the mechanism of the theoretical model in a specific case which can be described in empirical terms, and hence clarify the intuition behind the mechanism. We take a very simple case in which alternative natural resources are perfectly substitutable inputs in a Cobb–Douglas production function. Pollution is directly linked to the use of these natural resources, with some choices being cleaner than others; compare for instance coal and natural gas with respect to sulfur emissions. The social cost of using a natural resource is the sum of the private cost (i.e. the extraction cost) and the external cost (i.e. pollution damages). Since technological change is unbiased, the extraction cost remains constant. However the marginal damage cost of using a given input increases with income. Hence when income is low the extraction cost dominates the damage cost, and the cheapest input is chosen. However, when income is high the damage cost dominates the extraction cost, and the cleanest input is chosen. If there is a zero-emissions alternative it will dominate in the long run.

3.1 The environment

There is a unit mass of competitive firms which produce a single aggregate final good the price of which is normalized to 1. Both the firms and the population L are spread uniformly over a unit area of land. The utility function has the form of equation 4, $W = X/f(P)$, where $f(P) = \exp(P^{\phi})$:

$$W = X/\exp(P^{\phi}). \tag{9}$$

We thus have multiplicative utility, and in terms of equation 2 we have $\eta = 1$.

The representative firm in symmetric equilibrium hires an effective labour–capital aggregate A and buys a resource-intensive intermediate input R; that firm's production function is

$$Y(t) = A(t)^{1-\alpha} R(t)^{\alpha}, \tag{10}$$

where α is the share of the intermediate input, P is the aggregate flow of pollution—which is uniformly mixed—and ϕ is a parameter greater than 1. Effective labour–capital A grows at a constant rate g: $\dot{A}(t)/A(t) = g$. From now on we omit the time index whenever possible.

The intermediate input R—which we can think of as electricity—is the sum of inputs from n different resource-based technologies, which are all perfect substitutes in production. The quantity of input from technology j is denoted D_j, so

$$R = \sum_{j=1}^{n} D_j. \tag{11}$$

moves rapidly to the left, i.e. towards lower emissions. On the other hand, when the optimal point is at a kink then there is a jump up in marginal abatement cost, and there is a period during which increases in income lead to no increases in abatement effort, hence polluting emissions P grow in line with A and X.

The use of input quantity D_j leads to emission of pollution $\psi_j D_j$, where $\psi_j \geq 0$, hence aggregate pollution

$$P = \sum_{j=1}^{n} \psi_j D_j. \tag{12}$$

The cost of a unit of input j is w_j.

We can interpret alternative technologies j and k simply as alternative resource inputs, for instance low- and high-sulfur coal for electricity generation. However, a third technology l could be high-sulfur coal combined with flue-gas desulfurization (FGD). If the input is simply a natural resource then we can think of it as being extracted competitively from a large homogeneous stock, with each unit extracted requiring w_j units of final good as input. But for technology l the price w_l would be w_k plus the unit cost of FGD, and unit emissions ψ_l would be $\psi_k \times$ the fraction remaining after FGD.

Now consider just one technology j, and show (using equations 10–12) that the PPF for net final-good production X and the flow of pollution P if that technology is used exclusively is described by the following function:

$$X_j = A^{1-\alpha}(P/\psi_j)^{\alpha} - w_j P/\psi_j. \tag{13}$$

And if there is some technology for which $\psi_j = 0$ then this technology is *clean* (Definition 2), and we have (after solving the representative firm's optimization problem to decide the quantity of R_j to use) a single point

$$X_j = A(\alpha/w_j)^{\alpha/(1-\alpha)}(1-\alpha). \tag{14}$$

Since the natural-resource inputs are perfect substitutes, the overall PPF is simply the convex combination of the alternative PPFs defined above. In Figure 4(a) we illustrate the PPF (and one indifference curve) when there are three alternative technologies, two polluting and one clean.

3.2 The solution

In solving the model analytically we focus throughout on the social planner's solution; given this solution the regulatory problem is straightforward. To begin with assume just two technologies, so the planner chooses the set of values (D_1, D_2) to maximize W (equation 9). Take equation 9 and use equations (11–13) to get a version of the planner's problem:

$$\max_{D_1,D_2} W = [A^{1-\alpha}(D_1 + D_2)^{\alpha} - w_1 D_1 - w_2 D_2] \exp[-(\psi_1 D_1 + \psi_2 D_2)^{\phi}].$$

Now take the first-order conditions in D_1 and D_2 respectively to derive the following necessary conditions for an internal optimum:

$$\alpha Y/(D_1 + D_2) = w_1 + \phi\psi_1(\psi_1 D_1 + \psi_2 D_2)^{\phi-1}[Y - (w_1 D_1 + w_2 D_2)]; \tag{15}$$

$$\alpha Y/(D_1 + D_2) = w_2 + \phi\psi_2(\psi_1 D_1 + \psi_2 D_2)^{\phi-1}[Y - (w_1 D_1 + w_2 D_2)]. \tag{16}$$

In these equations, the marginal societal benefits of making an extra unit of intermediate good R ("electricity") using technology j are on the left-hand side, and the marginal costs are on the right-hand side. The marginal benefits are identical whether we use input 1 or 2 to make R, but the marginal costs differ. The costs are the sum of the natural-resource input costs w_j and the pollution damage costs.

To build intuition we start with the case in which $w_1 < w_2$ and $\psi_1 < \psi_2$, so D_1 is both cheaper and cleaner, and D_2 will never be used. Then we take the two-technology case, and finally multiple technologies.

Proposition 2

(i) *When only input D_1 is used, from any given initial state (defined by $A(0)$), P increases monotonically and approaches a limit of $\overline{P} = (\alpha/\phi)^{1/\phi}$. If we let $A(0)$ approach zero then the initial growth rate of P approaches g from below.*

(ii) *In a two-technology economy, there exist times T_{1a} and T_{1b} (where $T_{1b} > T_{1a}$) such that up to T_{1a}, D_1 increases monotonically while $D_2 = 0$. Between T_{1a} and T_{1b}, D_1 decreases monotonically while D_2 increases monotonically. And for $t \geq T_{1b}$, $D_1 = 0$ and D_2 increases monotonically. Furthermore, T_{1a} and T_{1b} can be expressed in closed form. In the special case of $\psi_2 = 0$ (the cleaner resource is perfectly clean) then T_{1b} is not defined; instead, as $t \to \infty$, $D_1 \to 0$, and hence $P \to 0$.*

(iii) *In an n-technology economy there is a series of m transitions (where $m \leq n-1$), starting with the cheapest input and ending with the cleanest. Each of these transitions proceeds analogously to the transition from technology 1 to 2. The remaining $n-m-1$ inputs are never used.*

Proof See Appendix A.

In Figure 4 we illustrate the development of the economy in a specific case with three technologies, the third of which is perfectly clean. Figure 4(a) shows the PPF, which is the convex hull of the PPFs for the three individual technologies. And in Figure 4(b) we show the path of pollution P, and the pollution limit $\overline{P}$. We also show—using dotted lines—the paths of P which would be followed if (respectively) only technologies 1 and 2 were available.

We start with the interpretation of the single-technology case. The shadow price of the polluting input to the social planner is the sum of extraction cost and marginal damages. The extraction cost is constant, whereas marginal damages increase in Y. So when Y is small the shadow price is approximately equal to the constant extraction cost, and both resource use and polluting emissions track growth. As Y increases, marginal damages increase and hence the shadow price of using the polluting input increases, braking the growth in its use. When Y is large marginal damages dominate the extraction cost, the shadow price of using the input grows at the overall growth rate, and emissions (and input use) are constant. So we have a transition from emissions tracking growth towards (in the limit) constant emissions. The link to the theoretical model is straightforward: utility is multiplicative ($\eta = 1$), and in the limit of high pollution costs (and hence negligible resource extraction costs) the PPF is Cobb–Douglas and $\sigma = 1$. Hence polluting emissions are—in the limit—constant (Proposition 1).

Now we take the more interesting case when technology 2 is more expensive but cleaner, i.e. $\psi_1 > \psi_2$. In this case, as Y increases, the increasing importance of pollution damages does not just lead to pollution abatement within technology 1—i.e. the substitution of labour–capital for D_1 in production—it also narrows the gap between the social costs of D_1 (cheap and dirty) and D_2 (expensive but cleaner). At some point the social costs are equal, and a transition to the cleaner technology begins. The link to the theoretical model (and Proposition 1) is again straightforward: as long as the cleanest technology is not perfectly clean we still have, in the limit, $\eta = 1$ and $\sigma = 1$, and emissions approach a constant

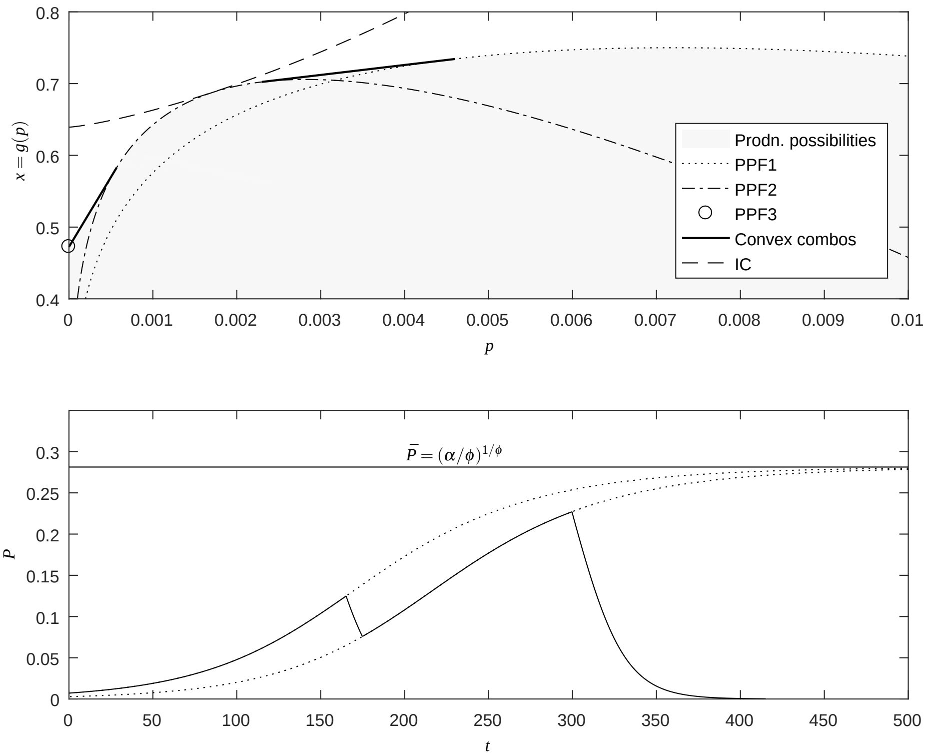

Fig. 4 (a) PPFs for three alternative technologies, one of which is clean (Definition 2), and the convex hull of those PPFs, which is the overall PPF. The PPF is invariant in A when plotted in (p,x) space. We also show a single indifference curve. (b) Pollution flow P compared to the limit, $\overline{P} = \alpha$. The dotted lines show pollution paths in case only one of the inputs is available. Parameters: $g = 0.02$; $A(0) = 1$; $\phi = 1.3$; $\psi_1 = 0.0072$, $\psi_2 = \psi_1/2$, $\psi_3 = 0$; $\alpha = 0.25$; $w_1 = \alpha$, $w_2 = 1.2\alpha$, $w_3 = 4\alpha$.

level in the long run. However, if the cleanest technology is perfectly clean then, in the limit, $\sigma \to \infty$ and polluting emissions approach zero.

In the multitechnology case, consider drawing the PPFs for each of the n technologies, as in Figure 4(a). The overall PPF is then the convex hull, and it is clear that if all the technologies are distinct then no more than two technologies will ever be used simultaneously, and that a subset of technologies will never be used at all because they are both expensive and dirty.

4 A calibrated model

The central hypothesis of the paper is that rising income drives the imposition of environmental regulations which—in the long run—drive switches to cleaner technologies and hence falling emissions. In this section we provide empirical support for this idea by showing that the timing of adoption of flue-gas desulfurization across six countries can be understood based on a model in which underlying preferences for clean air, and the unit cost of installing FGD, are constant across the countries and over time, and the timing of the imposition of the regulation is determined by income per capita, population, and the size of the territory.

In the introduction we argued that the shape of the PPF of pollution and production varies between countries, even those on the same income level, as does the shape of the indifference curves. Furthermore, biased technological change and new information may change PPFs and indifference curves over time.[7] It is therefore not possible to test the empirical relevance of the models above by looking for simple patterns such as turning points in pollution flows at given income levels. Instead of looking for patterns in emissions, we look for patterns in the application of environmental regulation, specifically the timing of adoption of FGD in Japan, the US, West Germany (as it was at the time of adoption), the UK, China, and India. FGD is a set of technologies used to remove sulfur dioxide from exhaust gases of coal-fired power plants (see US EPA (2003)). We choose it because of the readily available data about the timing of the implementation of FGD. We investigate the following hypothesis.

Hypothesis 1 *The unit costs of sulfur abatement through FGD are constant over time and across countries, and the time of introduction in a given country is determined by the marginal damage cost of sulfur emissions, which is a linear function of income per capita, and an increasing function of the size and population density of the country.*

We have data on the time of adoption (which we define as the first year when at least 5 percent of coal capacity has FGD installed), GDP per capita (from Maddison (2010)), population, and land area. The time of adoption ranges from 1970 (Japan) to 2016 (India).[8][9]

Ideally we would perform an econometric test of a structural model, but since we have only six observations we limit ourselves to a calibration exercise. We base the equation to be calibrated on equation 9, $W = X/\exp(P^{\phi})$. That is, we assume multiplicative utility following the climate literature. Since we do not have data on measured pollution concentrations, we assume that $\phi = 1$, making marginal damages approximately independent of P as long as total damages are small in relation to total utility. This assumption is also in line with the literature on damages from SO_2 where log-linear damages are typically assumed.[10] We then approximate X by real GDP, which we denote Y, and convert to per capita terms (so w is per

[7] For a specific example of the kind of idiosyncracies that may be relevant, consider sulfur emissions to the atmosphere in the UK and the US. In the UK there has been a rapid decline in SO_2 emissions since 1960, driven mainly by the replacement of coal by oil and gas in the overall energy mix. This shift was partly driven by the increase in road transport, but also by the ‘dash for gas’ in electricity generation, driven in turn by a steep decline in the price of gas relative to coal. In the US, sulfur emissions started to decline in the mid-1970s (see for instance Stern (2005)), at least partly due to the introduction of the clean air act in 1970. However, Ellerman and Montero (1998) demonstrate that the steep decline in sulfur emissions was facilitated by the significant fall in transport costs of coal which occurred subsequent to the deregulation of the railroads in the 1980s, which reduced the cost of shipping coal from the Powder River Basin; this coal is both the cheapest and cleanest in the US.

[8] , where the values for Germany are adjusted upwards by 14 percent to reflect the difference between average German GDP and West German GDP

[9] The year of FGD introduction is taken as the first year when at least 5 percent of coal capacity has FGD installed. The sources are as follows: Maxwell et al. (1978), Figure 2; US EPA (1995), Figure 4; Taylor et al. (2005) Figure 4; Markusson (2012) Table 1 (we assume that the 5 percent threshold was reached in 1993); Wang and Hao (2012), where the text implies that implementation of FGD took off around 2005; and lastly for India, Black and Veatch (2016), one of many available documents showing that India announced a stringent FGD program to start in 2016. GDP data is taken from Maddison (2010), extrapolated for India using equivalent data from the World Bank.

[10] See for instance Muller and Mendelsohn (2007), especially equation 12 in the additional materials, and the dose–response function of Barreca et al. (2017).

capita utility, and y per capita GDP):

$$w = ye^{-P}.$$

The next step is to think carefully about the implications of modelling different countries, which differ in surface area and population as well as GDP and polluting emissions. The concentration of pollution will (if the pollution is uniformly mixing and remains exclusively over the territory in question) be linearly related to emissions per unit of area, and damages (if they affect humans directly) should be a function of concentration. Denoting the area as H (recall that we previously normalized it to 1) we have

$$w = y \exp\left(-\frac{P/L}{H/L}\right).$$

This equation puts issues of *scale* into focus: it implies that if we replicate the economy (doubling P, L, and H but holding w and y constant) then the proportion of gross product y lost to pollution damages will remain the same. However, when we consider pollution transport it is clear that this will not in reality be the case: for an airborne pollutant, given a larger territory, a bigger proportion of emissions will land within the territory and thus cause damage there.

To account for pollution transport, we introduce a transport coefficient δ, where δ is the proportion of emissions transported out of the territory, and

$$\delta = \exp(-\theta H^{1/2}),$$

where θ is a positive parameter.[11] As $H \to 0$, $\delta \to 1$, and as $H \to \infty$, $\delta \to 0$, so for a very small territory almost all the pollution emitted leaves the territory without causing damage 'at home', whereas for a very large territory the reverse applies. So given δ we now have

$$w = y \exp\left(-(1-\delta)\frac{P/L}{H/L}\right).$$

Finally, and also related to scale, the above equation shows that when land area H increases, pollution damages decrease because the concentration of pollutant decreases. This effect should be straightforward if population and emissions are spread homogeneously over the territory. However, in reality they are spread inhomogeneously, and furthermore if the degree of inhomogeneity is an increasing function of the sparseness of population (because people concentrate in cities even in sparsely populated countries) then the effect of increasing H/L will be weakened. To allow for this possibility we introduce a parameter ω as follows:

$$w = y \exp\left(-(1-\delta)\frac{P/L}{(H/L)^{\omega}}\right).$$

So when $\omega = 1$ population is uniformly distributed, whereas when $\omega = 0$ overall population density has no effect because the population and electricity production are always confined to a sub-area in proportion to the size of the population. It remains to find marginal abatement benefits by differentiating wL w.r.t. P to obtain (after approximating $y = w$)

$$MAB = \phi(1-\delta)(L/H)^{\omega} y. \tag{17}$$

[11] The power of $1/2$ follows because if the area of the territory doubles, the average distance to the border is multiplied by $\sqrt{2}$.

To calibrate the model we must find values for θ and ω. We choose θ to match the observation of Smith and Jeffrey (1975) that around 75 percent of UK emissions leave the territory, yielding $\theta = 0.826$, and implying that in the largest countries (the US and China) around 83 percent of emissions cause damage within the territory. This leaves us with ω, which we choose in order to fit the data as well as possible, i.e. we find the value of ω which yields the set of six estimates for *MAB* with the lowest variance. This yields $\omega = 0.524$, implying that a doubling in population density leads to an increase in marginal abatement benefits by a factor of approximately $\sqrt{2}$.

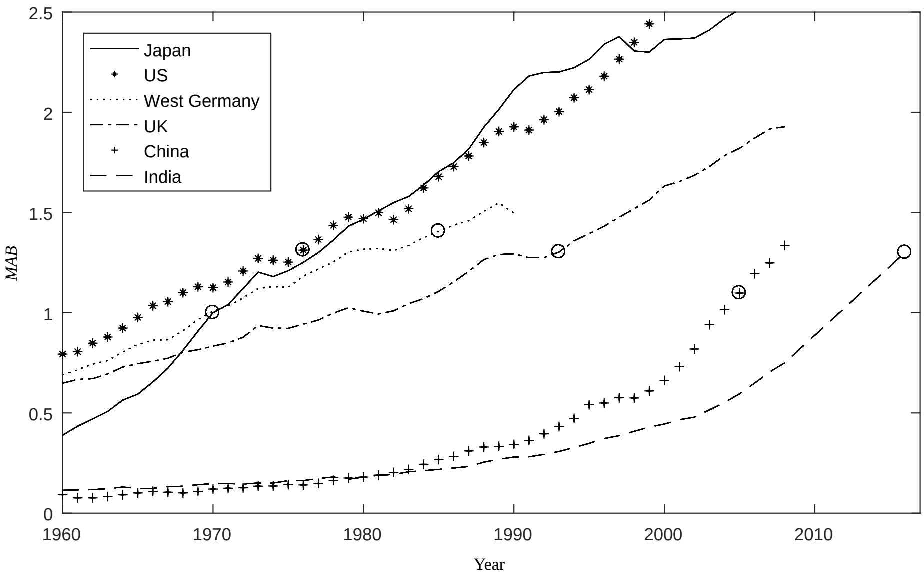

Fig. 5 Estimated marginal abatement benefits for the six countries plotted over time, and at the time of FGD adoption.

The results are illustrated in Figure 5, in which we show estimated marginal abatement benefits over time for each country, with a circle showing the time of FGD adoption. The cross-country variation in estimated *MAB* at the time of adoption gives an idea of the variation which is unexplained by the model. Note that—with the minor exception of Japan, which adopts 'early'—the countries adopt in the expected sequence and at expected times; small shifts in timing (between 0 and 3 years) would have all the other 5 countries adopting at the same level of estimated *MAB*. According to the estimates, both Japan and China adopt at somewhat lower benefit levels than the other four countries. These are also the two countries with the steepest rises in benefits of adoption, linked to their very high rates of economic growth at the time of adoption. In Japan this rapid growth—in both GDP and pollution flows—led to a dramatic increase in pressure for environmental improvements from the population, and the so-called 'pollution diet' of 1970; see Avenell (2012).

Figure 5 shows that we can rationalize most of the large differences in the time of adoption of FGD based on the model. Furthermore, inspection of the data shows that some of the simpler explanations that might be proposed are decisively rejected. For instance, there is no single level of GDP at which countries adopt FGD and thus reduce sulfur emissions. Furthermore, there is little evidence from the model that the unit costs of FGD have declined over time, thus encouraging lower-income countries to adopt at lower marginal benefit levels than the early-adopting higher-income countries.[12]

5 Previous literature

In this section we link our analysis more closely to existing theoretical literature. The discussion of theory starts with Grossman and Krueger (1991), who note that if pollution flows per unit of production decline, this decline must be the result of one or both of the following: (i) a change in the composition of consumption towards less polluting products; and (ii) a change in the choice of production technology used to produce the given product mix. Furthermore, they speculate about what economic forces might drive composition and technique effects, discussing the effects of trade and comparative advantage on composition, the effect of biased technological progress on the choice of technology, and finally the effect of growth on environmental policy (which could affect both the composition of consumption and the choice of technology). We now discuss these three potential drivers.

Regarding composition effects, Grossman and Krueger point out that if trade lies behind the local observations — perhaps because polluting firms moves out of countries in which they are strictly regulated, changing the composition of production locally — this strongly suggests that the declining trend will not hold globally. (This is linked to the pollution haven hypothesis; see for instance Levinson and Taylor 2008.) However, the pollution cases mentioned in the introduction (especially lead, CFCs, and SO_2), where emissions reductions are patently linked to the introduction and successive toughening of policy regulations, clearly demonstrate that there is much more to the pattern of rise and fall than trade and comparative advantage. Regarding the overall composition of consumption, Hart (2018a) shows that there has been a shift towards energy-intensive forms of consumption, which are also intrinsically pollution intensive. Ceteris paribus this should drive increases in polluting emissions faster than GDP growth, however this effect is counterbalanced by increasing energy-efficiency in producing goods, and hence a bias in technological progress.

Some form of bias in technological progress — driving a technique effect — seems more promising as a general mechanism, and several subsequent authors take up this idea; we discuss Andreoni and Levinson (2001), Brock and Taylor (2010), and Smulders et al. (2011). According to Andreoni and Levinson (2001) there are increasing returns to abatement, hence there is effectively a bias in technological progress such that the total cost of abating all polluting emissions approaches zero as labour–capital grows without bound, so pollution approaches zero under very mild assumptions regarding the utility function. Why such a change should occur in general is not clear, and none of the cases mentioned above in which sustained falls in pollution flows have been observed — lead, CFCs and SO_2 — seems to fit the Andreoni and Levinson mechanism.[13] Brock and Taylor (2010) also assume biased technological change which drives emissions down, however in their model pollution falls

[12] Note that, *ceteris paribus*, technological progress is not expected to drive down FGD costs. Technological progress implies that more goods can be produced using given inputs, however if it is neutral or unbiased then it will not change the relative prices of these goods. So in an economy with just two goods — an aggregate consumption good and sulfur capture through FGD — neutral technological progress implies that given inputs of labour–capital can produce more of both, but should not change the price of one relative to the other.

[13] For instance, we do not find any evidence that the cost of flue-gas desulfurization has declined significantly over time, nor that the scale of electricity production is a crucial factor. For a specific counterexample consider hydropower. At small scales, hydropower — both cheap and clean — may be available in sufficient quantity to meet demand. However, as the scale of the economy increases the marginal cost of hydro is likely to increase steeply, because of the limited flow of precipitation in a given geographical area.

monotonically on a balanced growth path. To generate the upward part of the curve they have to assume that the economy starts far from the balanced path, with far too little capital.[14] Brock and Taylor claim that their model matches aspects of the aggregate data, but there is no match between the model mechanism and the empirical cases such as those mentioned above. Smulders et al. (2011) assume that the scale effect is exactly cancelled by the underlying bias of technological change: growth is driven by the introduction of new technologies which require less labour per unit of production, and generate less pollution. However, there is a lag in learning about how to make each new technology clean, hence we have a series of rises and falls. The idea of lags driven by learning effects is clearly relevant empirically. However, Smulders et al. do not provide any guidance on the long-run trend in overall pollution: according to their model it is by assumption flat, but they could equally have assumed that it should be rising or falling.

Now we turn to the effect of growth on environmental policy. Grossman and Krueger claim (p.5) that 'more stringent pollution standards and stricter enforcement of existing laws may be a natural political response to economic growth.' This clearly fits the cases discussed above. In each case—lead, CFCs, SO_2 and NO_x, and CO_2—it can be argued that there is also an 'imperfect information' element to the pattern of technology adoption and subsequent regulation and clean-up. However, there is also a clear pattern of richer countries acting on the knowledge first in the case of local or regional pollutants such as lead, SO_2, and NO_x, and of richer countries leading the drive towards global regulation in the case of global pollutants such as CFCs and CO_2.[15]

To test the 'regulatory response to growth' mechanism we need a model without any of the other candidate mechanisms which might drive shifts in consumption patterns and technology choices, hence we need a model in which technological progress is neutral, and there is autarky, perfect information and optimal regulation. There are very few such models, and we discuss Figueroa and Pastén (2015) and Stokey (1998). Figueroa and Pastén (2015) treat pollution as an input in a CES production function together with effective labour–capital. The key to their model is the preference function, which is such that the price of pollution first rises slowly with increasing consumption, and later on more rapidly. Figueroa and Pastén argue that this function is intuitively reasonable because to the poor, pollution 'has the good smell of money' (p.92), whereas to the rich, it stinks. Again, although this idea may be reasonable in some cases, we claim that it lacks generality. Furthermore, our model demonstrates that such a preference structure is not necessary in order to deliver an EKC, once we model the production side correctly.

Stokey (1998) assumes separable preferences over consumption and pollution such that the elasticity of the price of pollution to the rate of consumption—$1/\eta$ in our notation, σ in Stokey's—is fixed. However, on the production side she assumes that, in effect, pollution is an input in a 'restricted' Cobb–Douglas production function. The restriction is that more polluting emissions boost production only up to a certain quantity, beyond which further emissions add nothing to production. This quantity increases linearly in effective labour–capital. Without the restriction—i.e. with a straightforward Cobb–Douglas production function—the trend in polluting emissions as labour–capital grows is monotonic, falling continuously if $1/\eta > 1$, rising continuously if $1/\eta < 1$. However, given the restriction we can obtain the

[14] Note that Ordás Criado et al. (2011) extend the Brock and Taylor model to allow for optimal policy, and in their model the pollution flow per capita is constant on a balanced growth path, hence there is no EKC.

[15] But note that the picture is complex for CO_2 due to factors including the great differences in the physical damages between countries (for instance, small island states with low levels of GDP per capita stand to suffer disproportionately), and strategic factors (countries with large reserves of oil or coal have an extra incentive to resist stringent global climate agreements).

hump-shaped path as long as $1/\eta > 1$ (i.e. the price of pollution is sufficiently sensitive to consumption): when capital is low the restriction binds and firms pollute at the maximum (which increases linearly over time), but at some point the restriction stops binding and pollution falls. The problem with the Stokey mechanism is, again, generality. Why should the relationship between polluting emissions and production have the assumed form? Why should the utility function have the assumed form? And how generally can we expect the condition of $1/\eta > 1$ to be fulfilled? Stokey does not address these questions, concentrating instead on a number of extensions to the basic model.

In our model we explicitly treat pollution as a by-product, thus going against a long tradition in theoretical work — going back at least as far as Baumol and Oates (1975) — of modelling firms' choice of polluting emissions by treating the flow of emissions as a freely disposable input; the more the firm emits, the more goods it can produce. Murty et al. (2012) challenge this tradition, arguing that pollution should be treated as a by-product of the use of natural resources. Furthermore, they show that in the field of DEA (data envelopment analysis), such a change of approach has a profound effect on the results of empirical analyses. The implications of the fact that pollution is a by-product have not been explicitly tackled in the literature on growth and pollution (but note the informal discussion of Smulders 2006). Some authors, such as Figueroa and Pastén (2015), simply treat pollution as a regular input. A more common approach — used by Stokey (1998), Andreoni and Levinson (2001), and Brock and Taylor (2010), following Copeland and Taylor (1994) — is to take the hybrid approach discussed in the previous paragraph.

In work focusing on specific pollutants — such as CO_2 — it is almost unavoidable to treat pollution as a by-product of natural resource use: see for instance Golosov et al. (2014) and Hart (2018b). The focus of Golosov et al. (2014) is the static price of carbon emissions, whereas Hart (2018b) performs a dynamic analysis of the balance between carbon pricing and research subsidies. In the latter paper we see how rising income pushes up the carbon price, which in turn drives a transition to clean technology. In other words, the key mechanism of the present paper is also operative in Hart (2018b).

6 Conclusions

The strongest prediction of the theoretical model is that if a clean technology exists, it should be only a matter of time — in an economy in which effective aggregate labour–capital grows without bound — before that technology is adopted. So the debate about optimal climate policy should be about the timing of a switch to clean technology rather than whether or not such a switch should be made. Furthermore, if incomes continue to increase, more and more substances will come to be considered pollutants and will be eliminated, as WTP for a pristine environment increases; consider for instance the current debates about microplastics.

The theoretical model is built on a very strong assumption about technological change — that it is unbiased — which we know does not hold in practice. The effect of adding directed technological change to the model would be to delay pollution reductions compared to the case where clean technologies are readily available without the need for research investments, but these reductions should be more abrupt (in a single-country context) and more coordinated (across countries).[16] A host of other potential extensions to the model would have similar effects, including allowing for stock effects, cross-border flows, and learning

[16] But note that in the case of FGD the calibrated model above suggests a limited role for DTC, and Hart (2013, 2018b) argues that the power of the DTC mechanism developed in models such as Acemoglu et al. (2012) is exaggerated when compared to reality.

about pollution damages. For instance, in the case of CFCs we have a stock pollutant which flows across borders, and where there was initially no knowledge of the damaging effects. The result was that when the effects were discovered, the rich countries were already well past the point at which they would have chosen (given perfect information) to halt emissions completely. They therefore did so abruptly. Furthermore, since the pollutant crosses borders they also ensured that countries with lower GDP—which would have preferred to carry on increasing their emissions at the time—followed suit; to induce them to do so they compensated the lower-income countries financially (see Sunstein 2007).[17] Inter alia, the CFC example shows that the oft-stated claim that the EKC applies to local but not global pollutants is wide of the mark. Through the lens of our model it should be clear that action on global pollutants may be delayed compared to a hypothetical case with a global government, due to the need for negotiations, free-riding incentives, etc. However—as the CFC case shows—where damage costs are sufficiently (and indisputably) large in relation to abatement costs, action is taken.

We postulated the CES utility function with very little discussion. Crucially, it implies that WTP for higher environmental quality Q approaches zero when income approaches zero, and approaches infinity when Q is bounded above and income approaches infinity. (These properties are all that are needed to generate the key results, the assumption of CES is made to rule out confounding mechanisms, similarly to the assumption of constant returns in the ppf.) Here we argue that these assumptions are very mild. It is hard to see how WTP for lower pollution flows P could fail to approach zero as long as $Q > 0$ and income approaches zero, and similarly it is hard to see how WTP for lower P could fail to approach infinity as long as Q is bounded above and income approaches infinity. However, there seems to be remarkably little research which systematically studies the WTP to reduce pollution or increase environmental quality as a function of income; for one example see Jacobsen and Hanley (2009). Regarding the exact specification of the utility function, the most common assumption in the non-EKC literature is that marginal damages from a given change in Q are proportional to GDP, i.e. multiplicative utility; see for instance climate models such as Nordhaus (2008) and Golosov et al. (2014), and the study of SO_2 policy of Finus and Tjotta (2003).

The model has profound implications for environmental policy. In an unequal world in which rich countries' citizens are concerned about the environmental damages caused by poorer countries, the model shows that promoting rapid income growth may yield better results than exerting pressure on those countries to sharpen their environmental regulations. And in a political arena of conflicting priorities the model shows that concerns about the environment are not a passing fad, but rather that ever-stricter environmental policies—and concomitant transitions to clean technologies—are inevitable, and this knowledge should permeate investment policy for both infrastructure and knowledge, and the current generation's attitude to irreversible damage to the natural environment: environmental regulations should not be sacrificed for the sake of growth, rather they should be strengthened in anticipation of future valuations, a result related to the Krutilla–Fisher–Porter model (see Porter 1982).

[17] Another factor which might lead to more abrupt and coordinated reductions in emissions would be a rise in the prices of natural resources such as coal the use of which leads to the by-production of pollution. However, there is little evidence for generalized resource scarcity driving up prices any time soon (see Hart and Spiro (2011) and Hart (2016)), and scarcity of specific resources may push pollution either way. For instance, natural gas (low sulfur) is likely to become increasingly scarce long before coal (high sulfur), and this will push the ppf to the right, tending to increase pollution.

Finally, the model focuses on pollutants in isolation (one at a time), and in this context it is easy to see how pollution-free production is possible. In practice multiple pollutants are frequently linked together. In many cases, a switch of technology will reduce several pollutants simultaneously, as when natural gas is used for electricity generation instead of coal, or when catalytic converters are added to car exhaust systems. However, there may also be trade-offs between pollutants, or more generally between different effects of economic activity on environmental quality. The ultimate trade-off may be over the use of the limited land area of the Earth: it may be used for economic activity, or reserved for nature, or the two may be combined. There is a trend in growing economies towards increasing areas being reserved for nature; for instance, species such as wolves which may interfere with economic activity, are being reintroduced or allowed to spread in Europe (see for instance Trouwborst 2010). Empirical and theoretical analysis focusing on long-run land allocation (rather than pollution flows) could be an important contribution to the debate about sustainability and growth.

A Proofs

Proof of Lemma 1

Regarding the initial limit, consider the PPF and indifference curves in (P,X) space. Assume $t = 0$, and let $A(0)$ decrease, approaching zero. The entire PPF then approaches the origin. Hence from the properties of the indifference curves, the slope at the point of tangency must approach zero. And from the properties of the PPF, the point of tangency must approach $(\bar{P},\bar{X})$, which is $(\bar{p},\bar{x})$ in (p,x) space.

Now let t (and hence also A) increase without bound. Consider a point of tangency between an indifference curve and the PPF in (P,X) space when $A = A^*$. Denote this point (P^*,X^*), and the slope of the tangent as m^*. Now let $A = sA$, where $s > 1$, and consider the point (sP^*,sX^*). This point lies on the new PPF, and the slope of the PPF at (sP^*,sX^*) is the same as at (P^*,X^*). However, the slope of indifference curve at (sP^*,sX^*) is greater than m^*, and the new point of tangency must lie to the left. Switching to (p,x) space this shows that the point of tangency moves to the left along $g(p)$ as A increases.

Finally consider the final limit. Take any point on the PPF in (p,x) space with strictly positive p, and let $A \to \infty$ (so both P and X at this point approach infinity). From the properties of the indifference curves, the slope of the indifference curve through this point must approach infinity, implying that the optimal point must (in finite time) move to the left of this point. Hence $\lim_{A\to\infty} p = 0$.

Proof of Lemma 2

Part (i) follows from Lemma 1, which tells us that $(p,x) \to (\bar{p},\bar{x})$, and from Definition 1, which shows that when $g'(p) = 0$, $\sigma = 0$.

To prove part (ii), assume that the relevant section of the function $g(p)$ (i.e. that between $(0,\underline{x})$ and $(\bar{p},\bar{x})$) consists of a series of n straight lines of decreasing gradient, joined to each other and the continuation of $g(p)$ beyond $(\bar{p},\bar{x})$ at n kinks. Along each straight segment, $g''(p) = 0$ hence σ is infinite, whereas at each kink $g''(p)$ is infinite and $\sigma = 0$.

To prove part (iii), note that when a clean technology exists the final section of $g(p)$ (closest to the $p = 0$ axis) is a straight line of positive gradient, as must be the final section of $G(A,P)$. Choose some A^* such that optimal pollution flows are strictly positive; now choose any value of P, denoted $P^\dagger$, such that the point $(A^\dagger,P^\dagger)$ is on the final (straight) section of $G(A,P)$, while $P^\dagger < P^*$. Within finite time the point of tangency must move to the left of the chosen point, hence as A increases without bound, $P \to 0$.

Proof of Lemma 3

In the first stage of the proof we use the implicit function theorem to derive two expressions for $\mathrm{d}P/\mathrm{d}A$, one in the case of $\eta \neq 1$, and one in the case of $\eta = 1$:

When $\eta \neq 1$ and the utility function is separable, we can use equations 1 and 3 to show that $W = v(X) - h(P) = v[G(A,P)] - h(P)$ and hence at an internal optimum

$$v'(X)G'_P = h'(P),$$

and the solution for X and P, given A, must satisfy the following two equations, where the first is the production function and the second is the optimality condition above:

$$G_1(A,X,P) = X - G(A,P) = 0;$$
$$G_2(A,X,P) = v'(X)G'_P - h'(P) = 0.$$

Using the implicit function theorem we can write

$$\begin{pmatrix} \frac{\mathrm{d}X}{\mathrm{d}A} \\ \frac{\mathrm{d}P}{\mathrm{d}A} \end{pmatrix} = -\begin{pmatrix} \frac{\partial G_1}{\partial X} & \frac{\partial G_1}{\partial P} \\ \frac{\partial G_2}{\partial X} & \frac{\partial G_2}{\partial P} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \frac{\partial G_1}{\partial A} \\ \frac{\partial G_2}{\partial A} \end{pmatrix}.$$

Perform the calculus and invert the matrix to yield

$$\begin{pmatrix} \frac{\mathrm{d}X}{\mathrm{d}A} \\ \frac{\mathrm{d}P}{\mathrm{d}A} \end{pmatrix} = \frac{1}{h''(P) - v''(X)(G'_P)^2 - v'(X)G''(P)} \begin{pmatrix} v'(X)G''(P) - h''(P) & G'_P \\ -v''(X)G'_P & 1 \end{pmatrix} \cdot \begin{pmatrix} -G'_A \\ v'(X)G''_{AP} \end{pmatrix}.$$

Finally use the expressions for η (equation 5, Remark 1) and σ (equation 7, Definition 1) to derive equation 18 (below) for $\mathrm{d}P/\mathrm{d}A$. (Note that the expression for $\mathrm{d}X/\mathrm{d}A$ shows that X is unambiguously increasing.)

Now we turn to the case of multiplicative utility, $W = X/f(P)$ (equation 4). Analogously to the above equations we have

$$G'_P = Xf'(P)/f(P),$$
$$G_1(A,X,P) = X - G(A,P) = 0,$$
and
$$G_2(A,X,P) = G'_P - Xf'(P)/f(P) = 0.$$

Follow a process precisely analogous to the above to obtain equation 19.

When $\eta \neq 1$,
$$\frac{\mathrm{d}P}{\mathrm{d}A} = \frac{v'(X)G'_AG'_P/G}{h''(P) - v''(X)(G'_P)^2 - v'(X)G''(P)}\left(-\frac{1}{\eta} + \frac{1}{\sigma}\right); \tag{18}$$
and when $\eta = 1$,
$$\frac{\mathrm{d}P}{\mathrm{d}A} = \frac{G'_A/G}{-G''_P/G'_P + f''/f'}\left(-\frac{1}{\eta} + \frac{1}{\sigma}\right). \tag{19}$$

Given the signs of the derivatives (which follow from the properties assumed of the utility function and the ppf), in both cases the sign of $\mathrm{d}P/\mathrm{d}A$ depends on whether $\sigma \lesseqqgtr \eta$: P is increasing when $\sigma < \eta$, decreasing when $\sigma > \eta$, and constant when $\sigma = \eta$.

Proof of Proposition 2

(i) Take the FOC in D_1 (equation 15), set $D_2 = 0$ and insert $Y = A^{1-\alpha}D_1^\alpha$ and $P = \psi_1 D_1$:

$$\alpha A^{1-\alpha}D_1^\alpha = w_1D_1 + \phi(\psi_1D_1)^\phi(A^{1-\alpha}D_1^\alpha - w_1D_1). \tag{20}$$

Then apply the limits on A to derive expressions for P in the limit, and hence also the initial growth rate.

(ii) Up to some time T_{1a}, input 1 is used exclusively, and the quantity D_1 is the unique solution to equation (20). However, at T_{1a} the FOC in D_2 also holds (although $D_2 = 0$). We thus have two equations in D_1 and A. Use these to derive the expression for $D_1(T_{1a})$, and reinsert this expression into equation (20) to obtain an expression for $A(T_{1a})$, and finally use $A = A(0)e^{gt}$ to find T_{1a}. Finally use an equivalent procedure setting $D_1 = 0$ to derive the symmetric expression for $D_2(T_{2a})$ and T_{1b}.

$$T_{1a} = \frac{1}{g} \log \left[\frac{D_1(T_{1a})}{A(0)} \left(w_1 \frac{1 - \phi(\psi_1 D_1(T_{1a}))^\phi}{\alpha - \phi(\psi_1 D_1(T_{1a}))^\phi} \right)^{1/(1-\alpha)} \right], \tag{21}$$

and

$$T_{1b} = \frac{1}{g} \log \left[\frac{D_2(T_{1b})}{A(0)} \left(w_2 \frac{1 - \phi(\psi_2 D_2(T_{1b}))^\phi}{\alpha - \phi(\psi_2 D_2(T_{1b}))^\phi} \right)^{1/(1-\alpha)} \right], \tag{22}$$

where

$$D_1(T_{1a}) = \frac{1}{\psi_1} \left(\frac{\alpha \psi_1}{\phi} \frac{w_2 - w_1}{w_2 \psi_1 - w_1 \psi_2 - \alpha w_1 (\psi_1 - \psi_2)} \right)^{1/\phi} \tag{23}$$

and

$$D_2(T_{1b}) = \frac{1}{\psi_2} \left(\frac{\alpha \psi_2}{\phi} \frac{w_2 - w_1}{w_2 \psi_1 - w_1 \psi_2 - \alpha w_2 (\psi_1 - \psi_2)} \right)^{1/\phi}. \tag{24}$$

Between these limits we know that D_1 falls monotonically and D_2 increases, because (Lemma 1) the solution moves left along $g(p)$, hence it moves (monotonically) left along the set of convex combinations of the two technologies. As $\psi_2 \to 0$, $D_2(T_{1b}) \to \infty$, hence $T_{1b} \to \infty$, hence as $t \to \infty$, $D_1 \to 0$, and hence $P \to 0$.

(iii) The proof is straightforward, based on the heuristic explanation in Section 3.2, and left to the reader.

References

Acemoglu, D., Aghion, P., Bursztyn, L., Hemous, D., 2012. The environment and directed technical change. American Economic Review 102, 131–166.

Andreoni, J., Levinson, A., 2001. The simple analytics of the environmental Kuznets curve. Journal of Public Economics 80, 269–286.

Avenell, S., 2012. From fearsome pollution to Fukushima: Environmental activism and the nuclear blind spot in contemporary Japan. Environmental History 17 (2), 244–276.

Barreca, A. I., Neidell, M., Sanders, N. J., June 2017. Long-run pollution exposure and adult mortality: Evidence from the Acid Rain Program. Working Paper 23524, National Bureau of Economic Research.

Baumol, W. J., Oates, W. E., 1975. The theory of environmental policy. Prentice-Hall, New Jersey.

Black and Veatch, 2016. Air quality control solutions for India: Innovative approaches and seamless execution. Tech. rep.

Brock, W. A., Taylor, M. S., 2010. The green Solow model. Journal of Economic Growth 15 (2), 127–153.

Copeland, B. R., Taylor, M. S., 1994. North–South trade and the environment. The Quarterly Journal of Economics 109 (3), 755–787.

Ellerman, A., Joskow, P., Schmalensee, R., Bailey, E., Montero, J., 2000. Markets for Clean Air: The U.S. Acid Rain Program. Cambridge University Press.

Ellerman, A. D., Montero, J.-P., 1998. The declining trend in sulfur dioxide emissions: Implications for allowance prices. Journal of Environmental Economics and Management 36 (1), 26–45.

Figueroa, E., Pastén, R., 2015. Beyond additive preferences: Economic behavior and the income pollution path. Resource and Energy Economics 41, 91–102.

Finus, M., Tjotta, S., Sep. 2003. The Oslo Protocol on sulfur reduction: the great leap forward? Journal of Public Economics 87 (9-10), 2031–2048.

Golosov, M., Hassler, J., Krusell, P., Tsyvinski, A., 2014. Optimal taxes on fossil fuel in general equilibrium. Econometrica 82 (1), 41–88.

Grossman, G., Krueger, A. B., 1991. Environmental impacts of a North American free trade agreement. NBER Working Paper (No. 3914).

Grossman, G. M., Krueger, A. B., May 1995. Economic growth and the environment. Quarterly Journal of Economics 110 (2), 353–377.

Hart, R., 2013. Directed technological change and factor shares. Economics Letters 119, 77–80.

Hart, R., 2016. Non-renewable resources in the long run. Journal of Economic Dynamics and Control 71, 1–20.

Hart, R., 2018a. Rebound, directed technological change, and aggregate demand for energy. Journal of Environmental Economics and Management 89, 218–234.

Hart, R., 2018b. To everything there is a season: Carbon pricing, research subsidies, and the transition to fossil-free energy. Journal of the Association of Environmental and Resource Economists. Forthcoming.

Hart, R., Spiro, D., 2011. The elephant in Hotelling's room. Energy Policy 39 (12), 7834–7838.

Hicks, J., 1932. The Theory of Wages. Macmillan, London.

Jacobsen, J. B., Hanley, N., Jun. 2009. Are there income effects on global willingness to pay for biodiversity conservation? Environmental & Resource Economics 43 (2), 137–160.

Levinson, A., Taylor, M. S., 2008. Unmasking the pollution haven effect. International Economic Review 49 (1), 223–254.

Lopez, R., 1994. The environment as a factor of production: the effects of economic growth and trade liberalization. Journal of Environmental Economics and Management 27 (2), 163–184.

Maddison, A., 2010. Historical statistics of the world economy: 1–2008 ad. Tech. rep., Groningen growth and development centre.

Markusson, N., 2012. The politics of FGD deployment in the UK (1980s–2009). Tech. rep., UK ERC.

Maxwell, M., Elder, H., Morasky, T., 1978. Sulfur oxides control technology in Japan. Interagency task force report.

Muller, N. Z., Mendelsohn, R., 2007. Measuring the damages of air pollution in the United States. Journal of Environmental Economics and Management 54 (1), 1–14.

Murty, S., Russell, R. R., Levkoff, S. B., 2012. On modeling pollution-generating technologies. Journal of Environmental Economics and Management 64 (1), 117–135.

Nordhaus, W., 2008. A Question of Balance: Weighing the Options on Global Warming Policies. Yale University Press.

Ordás Criado, C., Valente, S., Stengos, T., 2011. Growth and pollution convergence: Theory and evidence. Journal of Environmental Economics and Management 62, 199–214.

Panayotou, T., 1993. Empirical tests and policy analysis of environmental degradation at different stages of economic development. ILO Working Papers 292778, International Labour Organization.

Porter, R. C., 1982. The new approach to wilderness preservation through benefit–cost analysis. Journal of Environmental Economics and Management 9 (1), 59–80.

Selden, T. M., Forrest, A. S., Lockhart, J. E., 1999. Analyzing the reductions in U.S. air pollution emissions: 1970 to 1990. Land Economics 75 (1), 1–21.

Smith, F., Jeffrey, G., 1975. Airborne transport of sulphur dioxide from the U.K. Atmospheric Environment 9, 643–659.

Smulders, S., 2006. Growth and environment: on U-curves without U-turns. In: de Miguel, C., Labandeira, X., Manzano, B. (Eds.), Economic modelling of climate change and energy policies. Edward Elgar, Cheltenham.

Smulders, S., Bretschger, L., Egli, H., 2011. Economic growth and the diffusion of clean technologies: Explaining environmental Kuznets curves. Environmental & Resource Economics 49 (1), 79–99.

Stern, D. I., 2004. The rise and fall of the environmental Kuznets curve. World Development 32 (8), 1419 – 1439.

Stern, D. I., 2005. Global sulfur emissions from 1850 to 2000. Chemosphere 58, 163–175.

Stern, N., 2008. The economics of climate change. American Economic Review 98 (2), 1–37.

Stokey, N. L., 1998. Are there limits to growth? International Economic Review 39, 1–31.

Sunstein, C. R., 2007. Of Montreal and Kyoto: A tale of two protocols. Harvard Environmental Law Review 31, 1–65.

Taylor, M. R., Rubin, E. S., Hounshell, D. A., 2005. Regulation as the mother of innovation: The case of SO_2 control. Law and Policy 27, 348–378.

Trouwborst, A., 2010. Managing the carnivore comeback: International and EU species protection law and the return of lynx, wolf and bear to Western Europe. Journal of Environmental Law 22 (3), 347–372.

US EPA, 1995. Flue gas desulfurization technologies for control of sulfur oxides: Research, development, and demonstration. Tech. rep.

US EPA, 2003. Air pollution control technology fact sheet. EPA-452/F-03-034.

von Storch, H., Costa-Cabral, M., Hagner, C., Feser, F., Pacyna, J., Pacyna, E., Kolb, S., 2003. Four decades of gasoline lead emissions and control policies in Europe: a retrospective assessment. Science of the Total Environment 311 (1), 151–176.

Wang, S., Hao, J., 2012. Air quality management in China: Issues, challenges, and options. Journal of Environmental Sciences 24 (1), 2–13.