Patent Protection and the Transition to Clean Technology*

Maria Alsina-Pujols† Isabel Hovdahl‡

August 16, 2023

Abstract

We analyze the use of patent protection as a new policy to direct technical change to clean technology. Contrary to popular belief, it is dirty (and not clean) innovations that should be excluded from patent protection to reduce emissions. In the short-run, removing patent protection on dirty technology increases emissions. However, the reduced markup on dirty technology can induce clean innovation, reducing emissions in the long-run. We use a general equilibrium model to show both analytically and numerically that removing patent protection on dirty technology can indeed promote the energy transition and reduce the cost of mitigating climate change.

Keywords: Patent protection, directed technical change, climate policy

JEL classification: H23, O33, O34, Q54, Q55, Q58

*We thank Lint Barrage, Simon Dietz, David Hémous, Knut-Einar Rosendahl, and Ragnar Torvik for their helpful comments. We also thank seminar and conference participants at ETH, LSE, UZH, YSEM 2023, Morzine Harnessing Innovation Workshop 2022, AERE 2021, EAERE 2021, EEA 2021, IEW 2021. All errors are our own.

†Department of Management, Technology and Economics, ETH Zürich. E-mail address: malsinapujols@ethz.ch

‡Department of Business and Management Science, NHH Norwegian School of Economics. E-mail address: isabel.hovdahl@nhh.no

1 Introduction

A broad literature has identified carbon taxes and research subsidies as the most efficient tools to induce innovation in carbon-free technologies and promote the energy transition (e.g., Bovenberg and Smulders, 1995, Popp, 2004, Fischer and Newell, 2008, Acemoglu et al., 2012, Bretschger et al., 2017). However, the policy rates recommended in the literature tend to exceed what most policy makers are willing to implement and the distortions associated with these policies are often ignored. In this context, this paper analyzes the use of patent protection as a new policy tool to help induce the transition to clean technology. The policy consists of removing patent protection on dirty technology that contributes to global warming. We explore this novel policy in a general equilibrium model with directed technical change and characterize under which conditions our patent policy is effective in inducing innovation in clean technology.

Patent protection has been accused of slowing down the energy transition by restricting access to technologies that can reduce emissions. A proponent of this view is Elon Musk, who, in 2014, issued a statement in which he released all of Tesla's patents to the public. He claimed that the patent system was partly to blame for the low market share of electric vehicles.[1] In recent years, some countries have even proposed to exclude clean technology from international patent agreements to promote the diffusion of clean technology.[2] This argument, however, ignores the role of patent protection in incentivizing investments in innovation.[3] We contribute to the literature by introducing and analyzing the novel use of patent protection as a new policy tool to induce clean innovation. Contrary to popular belief, we find that removing patent protection on *dirty*, and not clean technology, can in fact foster

[1]However, there is no free lunch. The patent pledge came with several caveats; most importantly, by using Tesla's patents, you cannot enforce patent rights against any party using your patented technology, including Tesla. https://www.tesla.com/blog/all-our-patent-are-belong-you

[2]https://www.wto.org/english/tratop_e/trips_e/cchange_e.htm

[3]Furthermore, it is not clear whether removing patent protection on clean innovations would in fact lead to increased diffusion of clean technology. See Hall and Helmers (2010) for an overview of existing evidence on the link between intellectual property rights and the development and transfer of clean technology.

the transition to clean technology and reduce the cost of mitigating climate change when traditional climate policy is restricted.

We build our model on Acemoglu et al. (2012) – henceforth, AABH. To analyze the role of patent policy and exclusivity rights on the energy transition, our model differs from that of AABH in some important aspects. First, we take into account the monopoly distortion generated by the innovation sector.[4] This change allows us to capture the total effect on the energy transition of removing patent protection on dirty technology.[5] Second, we assume that unsuccessful scientists do not have monopoly rights over old technologies. This limits the extent of monopoly distortion in the economy to a reasonable level. Third, to approximate the difficulties policy makers face when implementing high levels of carbon taxes and public funding of innovation, we introduce efficiency losses in transfers of public funds. Depending on the magnitude of these losses, financing climate policy can become very expensive.

The protection granted by patents can be reduced by limiting either patent length or breadth, i.e. the extent of the technological market in which the innovation has exclusivity.[6] In fact, the value of patent protection is to a large extent decided by patent breadth as it is the breadth that restricts patent holders' ability to exploit their market power.[7] In our model, we fix patent length and reduce patent breadth by removing patent protection on dirty innovations.[8] Since broader patents are equivalent to stronger patent protection, we use the terms "patent protection" and "patent breadth" interchangeably. As in AABH, successful

[4]It is common in the literature to introduce production subsidies to avoid the use of climate policy to correct for monopoly power in the innovation sector.

[5]It is important to highlight that our policy not only tackles an environmental failure, but it also corrects for a non-environmental issue, namely the monopoly distortion created by the patenting system.

[6]While the early literature on patent protection and growth focused on the optimal lifetime of patents, the literature has more recently emphasized the importance of patent breadth (e.g. Gilbert and Shapiro, 1990, Klemperer, 1990, Gallini, 1992).

[7]Patent length is not necessarily a binding constraint since a patent becomes obsolete once a new innovation arrives, which could explain why most patents elapse before their statutory length (see e.g. Lanjouw, Pakes, and Putnam, 1998, Arora, Ceccagnoli, and Cohen, 2008).

[8]We use the term "reduce" in this context because many innovations have several applications that include both dirty and clean technologies (and non-environmental applications). Removing patent protection on dirty innovations implies that the patent breadth of these patents is reduced to include only non-dirty applications.

innovators are granted a patent that protects them from competition for one period.[9] Only innovations in clean technology are granted patent protection.[10] We approximate the loss of patent protection on the private value of innovation by placing a cap on the markup that innovators can charge for their innovation in absence of a patent.[11]

The main challenge of our policy is that, unlike more standard policies, its effectiveness in inducing clean innovation is not straightforward. Removing patent protection on dirty technology has opposing effects on the profitability of clean innovation relative to dirty innovation, and hence on the transition to clean technology. On the one hand, our policy reduces the monopoly power of innovators in dirty technology, which reduces the expected profits of dirty innovation and induces innovation in the clean sector. On the other hand, the reduction in the price of dirty technology results in increased demand for dirty technology, which increases emissions in the short-run and can even push innovators back to the dirty sector. Another challenge of this policy is its implementation, which is discussed extensively in the final section of the paper.

We characterize under which conditions the net effect of our policy is such that it increases the relative profitability of clean innovation. The net effect depends on the elasticity of substitution between clean and dirty production. When clean and dirty inputs are "weak" substitutes, the price effect unambiguously dominates the market size effect. Intuitively, with a low elasticity of substitution, a reduction in the price of the dirty good will lead to only a modest increase in demand for the good. In that case, we show analytically that our policy is effective in inducing clean innovation. However, when the two inputs are strong substitutes, the increase in the market share of the dirty input could potentially be sufficient to induce

[9] One period in our model corresponds to 5 years. Therefore, we capture the relationship between patent protection and the private value of innovation by assuming that patents are broad but short-lived.

[10] However, the loss of patent protection does not result in complete removal of monopoly power since innovators can still protect their innovation through other channels such as secrecy or first-mover advantage. In fact, survey data has found that patents are not the main strategy for protecting innovations in many industries (e.g. Mansfield, 1986, Levin et al., 1987).

[11] This is a standard way of incorporate patent breadth in the literature (e.g. Li, 2001, Chu, 2009, Zeng, Zhang, and Fung, 2014).

innovation in the dirty sector.

To gain additional insight into the dynamic effect of our patent policy on the transition to clean technology, we perform a quantitative exercise in which we simulate the model for different assumptions regarding the private value of patent protection. First, we explore the short-run effect of our patent policy by simulating a static economy without climate policy. Our results show that removing patent protection on dirty technology does indeed increase emissions in the short term. However, it also increases the relative profitability of clean innovation for a wide range of reasonable parameter assumptions. To further explore the dynamic effect of our patent policy, we simulate the economy over time, with and without climate policy in place. Although our patent policy is not sufficient in itself to induce the transition to clean technology, it can still delay the climate catastrophe and, more importantly, it can substantially lower the cost of mitigating climate change when there are costs associated with traditional climate policy.

This paper relates to two strands of the literature on endogenous growth. The literature on directed technical change has emphasized the need for policies that increase the relative profitability of clean innovation (Acemoglu et al., 2012, 2016, Hémous, 2016, Greaker, Heggedal, and Rosendahl, 2018, Hart, 2019). Most papers have restricted their analysis to the use of carbon taxes and R&D subsidies, and has not explored alternative policy tools to induce clean innovation.[12] Another strand of the literature has investigated the relationship between patent policy and growth. Optimal patent protection must balance the dynamic gain from increased innovation with the static efficiency loss caused by monopoly power (Nordhaus, 1969, Kamien and Schwartz, 1974).[13] We contribute to the literature by exploring the use of patent policy to foster sustainable growth in light of climate change.

[12]However, some papers have considered potential distortions caused by environmental policy, e.g. Acemoglu et al. (2016) introduce an efficiency loss in the clean R&D subsidy and a carbon tax cap, while Hart (2019) combine an efficiency loss in the clean R&D subsidy with a deadweight loss in the carbon tax in the form of an enforcement cost.

[13]In general, this optimal trade-off can be achieved by granting patents that are either short-lived but broad, or narrow but long-lived (e.g. Klemperer, 1990, O'Donoghue, Scotchmer, and Thisse, 1998).

Although some papers have explored the effect of incomplete patent protection on optimal climate policy (e.g. Greaker and Pade, 2009, Greaker, Heggedal, and Rosendahl, 2018), few papers have considered the use of patent protection as a policy tool. One exception is Gerlagh, Kverndokk, and Rosendahl (2014), who studies the optimal time path of clean energy innovation in an expanding variety model. They find that when R&D subsidies are constrained to be constant, optimal investment in energy innovation can still be achieved by adjusting the length of patents.[14] We, on the other hand, study the use of a differentiated system of patent protection to induce clean technical change. Instead of directly incentivizing clean innovation, policy makers can induce clean innovation by decreasing the profitability of dirty innovation by limiting the breadth of patents on such innovations.[15] Using patent protection as a new policy tool to promote the energy transition offers the advantages that it does not require increased public expenditures and there is already substantial international collaboration among countries on intellectual property rights.

The remainder of the paper is organized as follows. Section 2 provides empirical evidence to motivate the need for additional efforts to induce clean innovation and the assumption of a relationship between patent protection and technology markup. Section 3 outlines the model used to analyze our patent policy. Section 4 explains the calibration of the model for the quantitative exercise, while Section 5 presents the main results. Section 6 tests the robustness of the main results for some extensions of the model. Section 7 concludes and provides a discussion of the feasibility of our patent policy and some potential challenges to excluding dirty technology from international patent legislation.

[14] In a similar vein of research, Suphaphiphat, Peretto, and Valente (2015) explores how different regimes of (physical) property rights on renewable resources affect innovation and sustainability, while Hori and Yamagami (2018) explores whether intellectual property rights can be used to alleviate resource scarcity and sustain growth when production requires the use of an exhaustible resource.

[15] In theory, one could also implement a tax on the profits from dirty patented innovations, although it is not straightforward how such a tax would be implemented.

2 Motivational evidence

In this section, we document two trends that motivate some of the key assumptions in this paper. First, despite the need to decarbonize energy production to reach net zero emissions, innovation efforts are still largely directed towards fossil fuel technologies. This motivates our assumption that additional policies are needed to induce innovation towards clean technologies. Second, we provide evidence on a relationship between firm profits and patent breadth. We find that patents with a smaller breadth are associated with lower firm profits, which motivates our assumption that removing patent protection reduces the price markup.

2.1 Innovation trends

We discuss the evolution of innovation efforts in clean and dirty technology by looking at patent applications, which is a common proxy for innovation trends (Griliches, 1990). Despite the importance of decarbonizing energy production, innovation in the energy sector is strongly biased towards dirty technologies. Figure 1 shows the evolution of dirty, clean, and gray electricity innovation in OECD countries, measured as the ratio of patent applications to total applications at the European Patent Office (EPO).[16] The figure shows that until 2005, dirty innovation was persistently higher than clean innovation. Then, clean innovation started to converge. However, after 2010, there has been a collapse in clean innovation, while dirty innovation has continued to increase.[17] The fact that dirty innovation still represents a large share of current innovation activity indicates the need for additional policies to help redirect innovation efforts.[18]

[16]A similar pattern is found for applications filled at the United States Patent Office (USPTO) (see e.g. Acemoglu et al., 2019).

[17]Acemoglu et al. (2019) find evidence that this decrease in green innovation could be explained by the shale gas revolution, which increased incentives to innovate in dirty technologies.

[18]Notice that this is not necessarily true for every sector. For example, in the transport sector, patent applications in clean transport technology have increased steadily since the 1990s, surpassing patent applications in dirty transport technology in the early 2000s (not shown). However, given that many clean

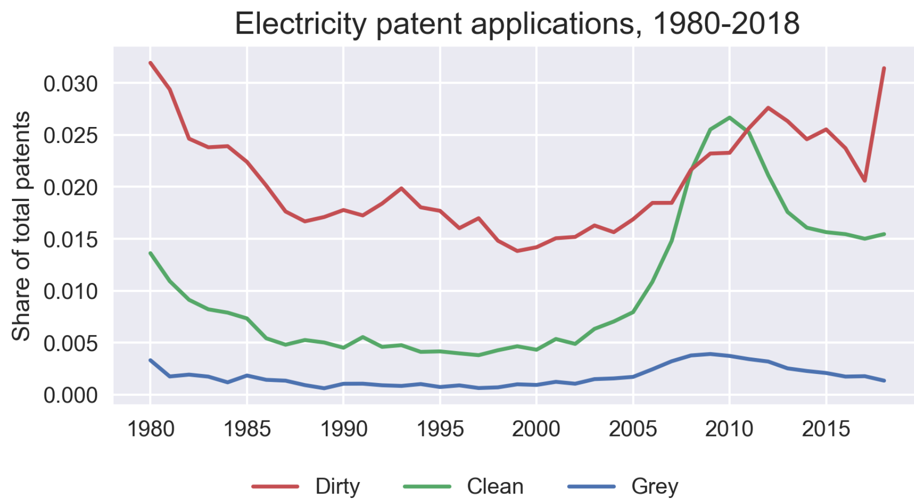

Figure 1: Based on data from PATSTAT. It includes all EPO electricity patents with an applicant or inventor from an OECD country. Patents are classified into grey, green and dirty following Dechezleprêtre, Martin, and Mohnen (2017), Table 21, which lists the CPC technological codes for each type. The authors identify grey codes as those that improve "the pollution efficiency of fossil technologies", thus, are harder to classify. When a patent has several types of CPC codes, we classify it into the type which has the higher number of codes. Draws are classified as grey.

2.2 Firm profits and patent breadth

Our theoretical model, detailed in the following section, features a novel policy that allows for differentiated levels of patent breadth depending on the innovation's nature. We assume that broader patents create more market power, allowing innovators to charge higher a markup and earn more profits. In this section, we provide some empirical evidence to justify our assumption of this direct relationship between patent breadth and firm profits.

Identifying the causal relationship between patent breadth and firm outcomes is challenging due to unobserved differences in firm quality. Although there is a large literature on the private value of patents, few papers have estimated the relationship between patent breadth and firm profits.[19] Therefore, we complement the literature by showing a strong relation-

transport technologies are powered by electricity, it remains vital to decarbonize electricity generation to reduce emissions in the transport sector.

[19] An early attempt was made by Lerner (1994), who found a positive correlation between the number of technology codes in a patent and firm profits in a cross-section of US firms. Hegde, Ljungqvist, and Raj

ship between patent breadth and firm profits for a panel of firms in multiple industries and countries, using different measures of patent breadth.

Our analysis combines data on firm profits from Compustat with data on the number of claims and technology codes in patent applications from PATSTAT. In order to match the two datasets, we use a third dataset provided by Bena et al. (2017)[20] that contains information on the firm's Compustat ID (gvkey) as well as on PATSTAT's patent number. Our sample includes firms from 45 countries during the period 1983–2019.

We use the number of claims and classification codes in a patent application to construct measures of patent breadth. The protection offered by a patent applies only to the technology described in the claims, and, therefore, it is generally preferred to write as many claims as possible in the application. Classification codes are used to describe the technological areas to which the invention in the patent contributes.[21] The codes are constructed so that the first digits represent a general technological field, while the following digits denote more specialized technology. In general, the more classification codes, the broader is the invention covered by the patent. Our measures of patent breadth are: i. "Claims": the number of claims made in a patent application; ii. "Codes Long": the number of CPC codes in their long format; iii. "Codes Medium": the number of CPC codes grouped by the so-called *main groups*, which typically have five- or six-digit CPC codes; and iv. "Codes Short": the number of four-digit CPC codes.

A simple regression of firms profits on patent breadth may suffer from strong endogeneity problems. For example, firms with higher profits may be able to innovate more and create broader innovations. In addition, unobserved quality differences between firms can affect patent breadth and firm profits at the same time (Hegde, Ljungqvist, and Raj, 2021). In an

(2021) restrict their analysis to US start-ups to exploit the quasi-random assignment of patent applications to examiners with different approval tolerance. They find causal evidence of a positive effect of broad patents on firm growth.

[20] See the Bena et al. (2017)'s Internet Appendix for a detailed description of the matching procedure.

[21] There are several classifications to define the technological areas and we use the Cooperative Patent International Patent Classification (CPC). Another widely used classification is the International Patent Classification (IPC). CPC is an extension of the IPC.

attempt to partially address these issues, we control for a rich set of fixed effects, at the firm and year level, as well as for the number of patents that a firm has in a given year.

To obtain a measure of the relationship between firm profits and patent breadth, we estimate the following equation

$$Profits_{ict} = \beta_0 + \beta_1 Breadth_{ict-1} + \beta_2 NumPatents_{ict-1} + \delta_i + \gamma_t + \epsilon_{ict}$$

where $Profits$ represent firm i, in country c, in year t. $Breadth$ is the average patent breadth, measured either as the average number of claims or CPC codes, while $NumPatents$ is the total number of patents for each firm i in year t. The independent variables are lagged by one year. δ and γ represent firm and year fixed effects, respectively.

Table 1 presents the regression results for our different measures of patent breadth. In column 3, the average number of claims is log-transformed, while in column 4, the dependent variable is a three-year moving average of firm profits. All specifications present a strong and statistically significant positive relationship. For example, column 2 shows that an additional claim is associated with an additional 0.604 million USD, while column 5 shows that an additional CPC code is associated with an additional 0.909 million USD. However, as shown in column 7, an additional broad technological area (short CPC code) assigned to the patent can be worth almost 11 million USD more in profits. Although our model specification does not entirely solve the endogeneity issue, it motivates our assumption of a positive association between patent protection and firm profits.

3 Theoretical framework

We consider a closed economy a la AABH. The four main differences between our model and the one in AABH are: i. the introduction of patent policy; ii. the presence of an uncorrected monopoly distortion in technology markets; iii. the fact that only successful innovators have

Table 1: Relationship between patent breadth and profits

	Profits						
	(1)	(2)	(3)	(4) MA3	(5)	(6)	(7)
Claims	0.419***	0.604***		0.665***			
	(0.09)	(0.15)		(0.15)			
Log(Claims)			186.836***				
			(43.23)				
CodesLong					0.909***		
					(0.20)		
CodesMedium						4.687***	
						(0.63)	
CodesShort							10.837***
							(1.89)
FE (Firm,Year)		√	√	√	√	√	√
Num. Patents	√	√	√	√	√	√	√
Observations	48974	48974	48974	44195	48974	48974	48974
Adj R^2	0.160	0.764	0.761	0.793	0.762	0.765	0.765
Countries	46	46	46	45	46	46	46

Note: Clustered standard errors on country-level in parentheses. Independent variable is lagged one year. Profits are measured in million USD. *** p<0.01, ** p<0.05, * p<0.1.

monopoly rights over new technologies; and iv. the consideration of policy inefficiencies (distortions).[22] We only briefly explain the main equations in the model and instead focus on the innovation sector and the role of patent policy.

3.1 Consumers and final good production

There is a representative household with a lifetime utility given by

$$U_0 = \sum_{t=0}^{\infty} \frac{1}{(1+\rho)^t} u(C_t) \tag{1}$$

[22]Other differences to AABH include intermediate input multiplicative-climate damages and decreasing returns to scientific labor.

where $u(C_t)$ is the instantaneous utility of consuming a unique final good at time t, and $\rho > 0$ is the discount rate. We assume that the instantaneous utility function increases in C_t ($u'(C_t) > 0$), is twice differentiable and concave ($u''(C_t) < 0$), and satisfies $\underset{C_t \to 0}{lim} u'(C_t) = \infty$. Time is discrete and runs to infinity.

Households consume a certain amount of a unique final good that is denoted by Y_t. Identical and perfectly competitive firms produce the final good by combining a clean and a dirty intermediate input, Y_{ct} and Y_{dt}, according to

$$Y_t = \left[Y_{ct}^{\frac{\epsilon-1}{\epsilon}} + Y_{dt}^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}}, \tag{2}$$

where ϵ is the constant elasticity of substitution between the inputs. Inputs are classified into clean or dirty depending on the technology they are produced with. For instance, in the energy sector, clean technologies are those based on renewable energy, while dirty technologies refer to the extraction and production of fossil fuel energy, for example, hydraulic fracturing.[23] We assume that the two inputs are gross substitutes, that is, $\epsilon > 1$, although the exact degree of substitutability is arguable.

3.2 Intermediate goods and the environment

Clean and dirty inputs are produced by identical and competitive firms that combine labor and a unit continuum of different machines, according to

$$Y_{jt} = \Omega(S_t) L_{jt}^{1-\alpha} \int_0^1 A_{ijt}^{1-\alpha} x_{ijt}^{\alpha} di, \tag{3}$$

where $\alpha \in (0, 1)$ and j indexes the sector, $j \in \{c, d\}$. Each machine is specific to a sector, with A_{ijt} being the quality of machine i used in sector j, and x_{ijt} being the quantity used of that machine. We assume drastic innovation, i.e., an innovation in machine i in sector j

[23] There is no gray technology in the model, i.e., energy-efficiency technology, nor is there any neutral technology.

causes the old vintage of the machine to be fully replaced by the new one. L_{jt} is the labor input, which is supplied inelastically. Normalizing the labor supply to unity, market clearing requires

$$L_{ct} + L_{dt} \leq 1. \tag{4}$$

We take the approach of climate damage being multiplicative to intermediate output.[24] $\Omega_t \in [0, 1]$ is the damage to production caused by an increase in atmospheric carbon concentrations, S_t, above pre-industrial levels. Under low levels of accumulated emissions, that is, low S_t, damage to production is minimal and $\Omega_t \to 1$. Accumulated emissions evolve according to a simple difference equation

$$S_{t+1} = \xi Y_{dt} + S_t, \tag{5}$$

where emissions from one unit of dirty production cause an increase in carbon concentrations by ξ units.[25] By assumption, S_{t+1} can only take values in the interval $(0, \bar{S})$. The upper bound on accumulated emissions, $\bar{S}$, captures the concern that there exists some tipping point that once reached, there is a climate catastrophe, in which case $\Omega_t = 0 \; \forall \; t$.

The intermediate good firm's problem is

$$\underset{L_{jt}, x_{ijt}}{max} \; p_{jt} Y_{jt} - w_t L_{jt} - \int_0^1 p_{ijt}(1 - z_j) x_{ijt} di \tag{6}$$

where p_{jt} denotes the price of the intermediate input of sector j, p_{ijt} is the price of each machine of type i in sector j, while z_j is a potential subsidy on the price of machines in sector j. The wage rate, w_t, must be the same across both sectors in equilibrium. The first-order

[24] In AABH it is trivial whether climate change reduces intermediate production or whether consumers have a dislike of climate change. However, this is no longer the case when there are efficiency losses associated with transfers of public funds. Due to efficiency losses, overall damages from climate change will be lower if climate change reduces utility as opposed to intermediate production.

[25] Notice that there could be some environmental regeneration over time caused by carbon sinks. However, a linear relationship between accumulated emissions and global warming has been found to be a good approximation of actual climate dynamics (Dietz and Venmans, 2019, Dietz et al., 2021).

condition for the optimal use of machine i results in the standard demand for the machine

$$x_{ijt} = \left(\frac{\alpha \Omega(S_t) p_{jt}}{p_{ijt}(1 - z_j)} \right)^{\frac{1}{1-\alpha}} A_{ijt} L_{jt}, \tag{7}$$

which is increasing in the level of technology, but decreasing in the price of the machine.

3.3 Supply of machines and patent policy

To produce one unit of a machine, ψ units of the final good must be used. There is perfect competition in machine lines without quality improvements. Since innovation efforts were unsuccessful, old vintages of these machines are instead sold at the competitive price,

$$p_{ijt}^{CO} = \psi. \tag{8}$$

In machine lines where innovation efforts have been successful, machines are patented and supplied by monopolists. Each monopolist maximizes profits according to

$$\max_{p_{ijt}, x_{ijt}} (p_{ijt} - \psi) x_{ijt}, \tag{9}$$

where demand for the machine, x_{ijt}, is given by equation 7. Maximizing profits results in the unconstrained monopoly price, which is a constant markup on the marginal cost of producing the machine, ψ/α. [26]

While most of the previous literature has (implicitly) assumed that successful innovators are able to fully protect their innovations, we remove access to patent protection on dirty innovations. Motivated by the literature, we implement this patent policy by placing a cap, μ, on the dirty markup. In this way, the monopolist of a dirty machine must charge a lower price than the unconstrained monopoly price. Under the new patent policy, the monopolist

[26] However, notice that the social planner can equate the price of machines to the competitive price by simply setting $z_j = 1 - \alpha$.

charges the following price

$$p_{ijt}^{MO} = \begin{cases} \frac{\psi}{\alpha} & if\, j = c \\ \mu\frac{\psi}{\alpha} & if\, j = d, \end{cases} \tag{10}$$

where μ is a value in the interval $(\alpha, 1]$.

If $\mu = 1$, dirty monopolists are granted full patent protection and can therefore charge the unconstrained monopoly price. If $\mu < 1$, dirty monopolists cannot register their innovations at the patent office and must therefore sell the machines at a lower price to discourage imitation. Notice that $\mu = \alpha$ implies that monopolists must sell their innovations at the competitive price when the innovation is not protected by a patent. We therefore assume that $\mu > \alpha$ since innovators have other ways, e.g. secrecy or first-mover advantage, to protect their innovation and generate some market power.[27] Importantly, the closer μ is to α, the more vital is patent protection to generate market power.

Combining the price of machines from equation 10 with the demand for machines from equation 7, the per-period profit of a producer of a clean machine becomes

$$\pi_{ict} = (1-\alpha)\left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{\alpha}{1-z_c}\right)^{\frac{1}{1-\alpha}}\left(\Omega(S_t)p_{ct}\right)^{\frac{1}{1-\alpha}}A_{ict}L_{ct}, \tag{11}$$

while the per-period profit of a producer of a dirty machine becomes

$$\pi_{idt} = (\mu-\alpha)\left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{\alpha}{\mu(1-z_d)}\right)^{\frac{1}{1-\alpha}}\left(\Omega(S_t)p_{dt}\right)^{\frac{1}{1-\alpha}}A_{idt}L_{dt}. \tag{12}$$

The cap on the price of dirty machines, μ, enters the expression for dirty profits twice. The first term, $(\mu - \alpha)$, captures the fact that dirty monopolists must now charge a lower price in the absence of patent protection, while the second term, $(\alpha/\mu(1 - z_d))^{1/(1-\alpha)}$, captures the fact that a reduced price will increase demand for dirty machines.[28]

[27] When $\mu < \alpha$, innovators must set the price below the marginal cost. We ignore this trivial case.

[28] Notice that this latter effect of the patent policy on dirty profits would disappear if the production subsidies are used to correct for the monopoly distortion. In that case, the required production subsidies

3.4 Innovation and allocation of scientists

Scientists choose to work in either the clean or the dirty sector. The total amount of scientists in each period is normalized to unity,

$$s_{ct} + s_{dt} \leq 1, \tag{13}$$

where s_{jt} denotes the mass of scientists working in sector j. Once scientists have chosen a sector they are randomly allocated to a specific machine in that sector.[29] Their scientific effort is successful with probability $\eta_j \in (0,1)$. If successful, the innovator becomes the monopolist producer of the latest version of that machine for one period.[30]

Successful innovations cause an increase in the productivity of the technology by a factor γ, and the new productivity level is given by $(1+\gamma)A_{ijt}$. The aggregate (and average) machine quality in sector A_{jt} is denoted by

$$A_{jt} = \int_0^1 A_{ijt} di. \tag{14}$$

Since scientists are unable to choose the specific machine, their decisions will be based on the average machine quality in each sector. Like Greaker, Heggedal, and Rosendahl (2018), we take into account the fact that more than one scientist might have the same successful innovation in a given period. This stepping-on-toes effect is represented by decreasing returns to scientific labor in each sector, s_{jt}^{σ}, with $\sigma \in (0,1)$. The probability of making a new innovation is therefore given by $\eta_j s_{jt}^{\sigma}$, resulting in the following evolution of the average

would be $z_c = 1-\alpha$ and $z_d = 1 - \frac{\alpha}{\mu}$.

[29]Alternatively, if we make the plausible assumptions that profits are increasing in the productivity of the machine, i.e. $\pi_{ijt} = A_{ijt}\left(p_{ijt} - \psi\right) x_{ijt}$, while the probability of innovating is decreasing in the productivity of the machine, i.e. $\frac{\eta_j}{A_{ijt}}$, scientists become indifferent with regards to which machine to choose. Mathematically, this would be equivalent to our assumption that scientists only choose the sector and not the specific machine.

[30]This assumption allows us to reduce the innovation problem to a static one in the decentralized economy. Greaker, Heggedal, and Rosendahl (2018) demonstrate the robustness of the model to a dynamic set-up where innovators remain the incumbent until replaced by entrants.

quality of machines in a sector,

$$A_{jt} = (1 + \gamma\eta_j s_{jt}^{\sigma})A_{jt-1}. \tag{15}$$

Using equations (11) and (12), and aggregating over the quality of machines in the sector, the expected profits of a scientist engaged in clean and dirty research are given by

$$\Pi_{ct} = \eta_c s_{ct}^{\sigma-1}(1-\alpha)\left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{\alpha}{1-z_c}\right)^{\frac{1}{1-\alpha}}(\Omega(S_t)p_{ct})^{\frac{1}{1-\alpha}}L_{ct}(1+\gamma)A_{ct-1} \tag{16}$$

and

$$\Pi_{idt} = \eta_d s_{dt}^{\sigma-1}(\mu-\alpha)\left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{\alpha}{\mu(1-z_d)}\right)^{\frac{1}{1-\alpha}}(\Omega(S_t)p_{dt})^{\frac{1}{1-\alpha}}L_{dt}(1+\gamma)A_{dt-1}, \tag{17}$$

where $\eta_j s_{jt}$ is the average productivity of a scientist entering sector j.[31] Technical change is driven by the relative profitability of clean research, which can be expressed as

$$\frac{\Pi_{ct}}{\Pi_{dt}} = \frac{\eta_c}{\eta_d}\left(\frac{s_{ct}}{s_{dt}}\right)^{\sigma-1} \times \underbrace{\frac{1-\alpha}{\mu-\alpha}\left(\frac{\mu(1-z_d)}{1-z_c}\right)^{\frac{1}{1-\alpha}}}_{\text{direct patent policy effect}} \times \underbrace{\left(\frac{p_{ct}}{p_{dt}}\right)^{\frac{1}{1-\alpha}}}_{\text{price effect}} \times \underbrace{\frac{L_{ct}}{L_{dt}}}_{\text{market size effect}} \times \underbrace{\frac{A_{ct-1}}{A_{dt-1}}}_{\text{direct productivity effect}}. \tag{18}$$

The direct effect of the patent policy (direct patent policy effect) is to reduce the price of dirty machines, which increases the relative profitability of innovating in clean technologies.

In addition to patent policy, technical change is driven by the standard channels in the literature by directing innovation towards the sector with the higher price (price effect), higher share of employment (market size effect) and the initially more advanced sector (direct productivity effect). Since removing patent protection on dirty innovations affects the relative price and labor share of the clean good, the policy will indirectly affect the relative

[31] Note that we divide the probability of success by the amount of scientists in a sector. Because we assume decreasing returns to the number of scientists in each sector, the solution for the allocation of scientists differs from the corner solutions found in AABH.

profitability of clean innovation through the price and market size channels as well.

The following proposition puts a sign on the direct and indirect effects of patent policy on the relative profitability of clean research:

Proposition 1. *All else equal, removing patent protection on dirty innovation, i.e., $\mu < 1$, has three partial effects on the relative profitability of clean research:*

i Direct effect of patent policy: In the absence of patent protection, successful innovators of dirty machines must sell at a lower price. Although they now face a higher demand for their machines, expected profits from dirty innovation fall, and thus the relative profitability of clean research increases.

ii Indirect effect through the price channel: Since dirty innovators must sell their machines at a lower price, the dirty input becomes less expensive. The relative price of the clean input increases, which increases the relative profitability of clean research.

iii Indirect effect through the market size channel: Assuming that the two inputs are substitutes, a lower price of dirty machines will lead to a re-allocation of labor to the dirty sector. The reduced market size of the clean sector will decrease the relative profitability of clean research.

See proof of Proposition 1 in Appendix B. Proposition 1 states the three effects of our patent policy on the relative profitability of clean innovation. Although two of the channels clearly induce innovation towards clean technology, the third channel works in the opposite direction. Hence, the net effect is ambiguous.

Inserting for the relative price and labor share of the clean good (see Appendix A), and assuming that there are no production subsidies to correct for the monopoly distortion (i.e., $z_j = 0$), the relative profitability of clean innovation can be expressed as a function of the share of clean scientists, average quality levels and policy instruments

$$\frac{\Pi_{ct}}{\Pi_{dt}} = (1+q_t)\frac{\eta_c}{\eta_d}\left(\frac{s_{ct}}{s_{dt}}\right)^{\sigma-1}\frac{1-\alpha}{\mu-\alpha}\mu^{\frac{1}{1-\alpha}}(1+\tau_t)^{\epsilon}\times$$
$$\left(\frac{\left(\eta_c s_{ct}^{\sigma}\left(\alpha^{\frac{\alpha}{1-\alpha}}-1\right)+1\right)(1+\gamma\eta_c s_{ct}^{\sigma})}{\left(\eta_d s_{dt}^{\sigma}\left(\left(\frac{\alpha}{\mu}\right)^{\frac{\alpha}{1-\alpha}}-1\right)+1\right)(1+\gamma\eta_d s_{dt}^{\sigma})}\right)^{-\varphi-1}\left(\frac{A_{ct-1}}{A_{dt-1}}\right)^{-\varphi}. \quad (19)$$

In addition to a carbon tax, τ_t, on the price of the dirty input, there is also a subsidy, q_t, given to scientists in the clean sector.

From equation 19 follows the next proposition on the net effect of removing patent protection:

Proposition 2. *Suppose that $1 \leq \epsilon \leq \frac{2-\alpha}{1-\alpha}$, i.e., the energy inputs are weak substitutes. Then, all else equal, our policy will lead to an increase of the relative profitability of clean research. When $\epsilon > \frac{2-\alpha}{1-\alpha}$, however, the effect of our policy on the relative profitability of clean research is still ambiguous.*

See proof of Proposition 2 in Appendix B. Proposition 2 states that for sufficiently low values of ϵ, that is, when the two inputs are weak substitutes, the price effect dominates the market size effect, and our patent policy unambiguously increases the relative profitability of clean innovation, inducing scientists to the clean sector. However, when the inputs are strong substitutes, the market size effect dominates the price effect, and it is ambiguous whether the direct effect of our patent policy dominates the net indirect effect.

Note that even though we cannot characterize it analytically, our patent policy can still incentivize clean innovation when ϵ is high. To gain further insight into how patent policy can help induce the energy transition, we next perform a quantitative analysis in which we simulate the model to explore the static and dynamic effects of removing patent protection on dirty technology at different values of ϵ.

4 Calibration

4.1 Optimization problem

For the numerical analysis, we assume a CRRA utility function

$$u(C_t) = \frac{C_t^{1-\nu}}{1-\nu}$$

The social planner's problem is to choose a path for the carbon tax, τ_t, and clean innovation subsidy, q_t, that maximizes discounted consumption. Both the research subsidy and the carbon tax are financed lump sum. However, we assume that there are inefficiencies associated with such government transfers. As a result, only a share $1 - d_q$ of the clean research subsidy and $1 - d_\tau$ of the carbon tax are refunded to consumers.

Consumption is equal to the production of the final good net of what is used up in producing machines and lost due to the efficiency losses in public transfers. Inserting for the budget restriction from equation 39 into the utility function, the maximization problem becomes

$$\max_{\tau_t, q_t} U = \sum_{t=0}^{T} \frac{1}{(1+\rho)^t} \frac{(Y_t - \psi(X_{ct} + X_{dt}) - d_q q_t \Pi_{ct} - d_\tau \tau_t p_{dt} Y_{dt})^{1-\nu}}{1-\nu}$$

In Appendix A, we show how Y_t, X_{ct}, X_{dt}, Π_{ct} and Y_{dt} evolve as functions of the technology levels, accumulated emissions and policy instruments. Additionally, technology levels evolve as functions of the share of clean scientists, and accumulated emissions evolve as a function of dirty production.

4.2 Parameter selection

We set a period in our model to 5 years, and we simulate the model for 80 periods when analyzing the long-run impacts of our patent policy. We follow AABH by setting $\nu = 2$

to match Nordhaus' intertemporal elasticity of substitution, the machine share of national income to $\alpha = 1/3$, the per annum probability of a successful innovation to $\eta_c = \eta_d = 0.02$, and the quality step to $\gamma = 1$. Without loss of generality, we normalize the cost of machines to $\psi = \alpha^2$. Although it is commonly accepted that there are decreasing returns in the innovation process, there is little empirical evidence on the exact magnitude of this effect. We follow Greaker, Heggedal, and Rosendahl (2018) and set $\sigma = 0.7$.[32]

To calibrate the efficiency loss associated with public funding of clean research, we compare the effectiveness of R&D in producing patents when the R&D is publicly or privately funded. See Section S.1 in Supplementary Material for the details of the estimation. We find publicly funded R&D to be 28% less effective in producing patents compared to privately financed R&D, and we set $d_q = 0.28$. For the efficiency loss in the carbon tax, we use the estimate from Barrage (2014) and set $d_\tau = 0.48$.

We initialize the simulations by computing the average quality levels of clean and dirty technology one period prior to the simulation runs, A_{c0} and A_{d0} (see equation 41 in Appendix A). Initial productivity levels are calculated using the global consumption of fossil and non-fossil fuel from BP (2022). From 2016 to 2020, the world consumed 464.26 and 2392.11 exajoules of energy from non-fossil and fossil energy sources, respectively.[33]

Cumulative emissions, S_t, evolve as a function of dirty production. An increase in atmospheric carbon concentrations causes an increase in the global mean temperature, resulting in economic damages, $\Omega(S_t)$. Thus, $\Omega(S_t)$ can be written as $\Omega(\Delta(T))$, where $\Delta(T)$ is the increase in the mean temperature above its pre-industrial level in degrees Celsius. The relationship between accumulated emissions, S_t, and global warming, $\Delta(T)$, can be approximated by a simple linear function (Dietz and Venmans, 2019, Dietz et al., 2021). We calibrate the damage function in AABH to correspond to a 10% decrease in production at $3°C$ warming

[32] Acemoglu et al. (2016) use a slightly lower estimate of 0.5 in their baseline analysis.

[33] To calculate the initial technology gap, we also need the initial share of clean scientists. If we set $s_{c0} = 1/3$ (and $\epsilon = 3$), this results in $A_{c0}/A_{d0} = 0.435$, that is, clean technology is initially 40% less productive than dirty technology. Note that the initial technology gap is not sensitive to the initial share of clean scientists. If we instead assume that $s_{c0} = 2/3$, we find that $A_{c0}/A_{d0} = 0.446$.

(e.g. Van Der Wijst et al., 2023). As in AABH, we assume that a climate catastrophe occurs at $6°C$ warming, in which case all production capabilities are lost, that is, $\Omega(6°C) = 0$. The calibration of the carbon cycle and the damage function is explained in detail in Section S.2 in Supplementary Material.

Although it is usually assumed that clean and dirty inputs are substitutes, the literature has used a wide range of estimates on the elasticity of substitution.[34] We use a mid-range estimate equal to $\epsilon = 3$, which is in line with micro-empirical estimates of the elasticity (e.g. Jo, 2020). However, we also check the robustness of our policy to lower and higher elasticity of substitution. For the discount rate, we use Nordhaus' preferred choice of 1.5% per annum discounting and set $\rho = 0.015$.

In Section 2, we presented some empirical evidence on the relationship between patent breadth and firm profits. However, given the lack of precise estimates of μ, i.e., the effect of patent protection on the price markup, we simulate the economy for a wide range of $\mu \in (\alpha, 1]$.

5 Main results

The following presents the main results of the quantitative analysis. We proceed in three steps. First, we compute the short-run effects on production and innovation of removing patent protection from dirty technology. Second, we explore the long-run effects by simulating the economy over time to see whether patent policy can induce the transition to clean technology. Third, we explore the policy impacts of our patent policy by simulating the economy with climate policy in place to see how patent policy affects the optimal carbon tax and R&D subsidy.

[34] Karydas and Zhang (2019) use a conservative elasticity of 0.7, while Greaker, Heggedal, and Rosendahl (2018) use an elasticity of 1.5 in their low substitution scenario. AABH use an elasticity of 3 in their low substitution scenario and an elasticity of 10 in their strong substitution scenario.

5.1 Short-run effects

We explore the short-run effects of our patent policy by calculating the static effect of a reduction in μ in period 0 before our simulation runs. Since there is no climate policy in place, there are also no losses associated with transfers of public funds. Table 2 shows the static change in production and relative profitability of clean innovation when $\mu < 1$ compared to when there is no patent policy ($\mu = 1$) for different assumptions regarding key parameters in the model. The table illustrates the two opposing effects of our patent policy in the short-run. Removing patent protection on dirty technology increases the production of dirty inputs, Y_d, and reduces the production of clean inputs, Y_c, which will lead to more emissions in the short-run. However, there is also an increase in the relative profitability of clean innovation, Π_c/Π_d, which can induce scientists to the clean sector, potentially reducing emissions in the long-run.

Table 2: Short-run effects of patent policy.

	$\%\Delta Y_d$	$\%\Delta Y_c$	$\%\Delta Y$	$\%\Delta \Pi_c/\Pi_d$
Baseline parameters				
$\mu = 0.8$	0.7%	-0.4%	0.4%	2.0%
$\mu = 0.6$	1.6%	-1.0%	1.0%	15.7%
$\mu = 0.4$	3.3%	-1.9%	2.0%	150.8%
Lower $\epsilon = 1.5$				
$\mu = 0.6$	1.3%	0.0%	0.8%	17.2%
Higher $\epsilon = 10$				
$\mu = 0.6$	2.7%	-5.9%	1.1%	8.9%
Higher $\eta_c = \eta_d = 0.2$				
$\mu = 0.6$	23.3%	-12.3%	14.0%	9.8%
Higher $\eta_c = \eta_d = 0.4$				
$\mu = 0.6$	87.8%	-34.7%	53.2%	-2.6%

In general, the static effects of our patent policy on production and innovation are increasing in the effect of our patent policy on the dirty mar-up. For example, when the price of dirty technology is limited to 40% of the monopoly markup in the absence of patent protection, that is, $\mu = 0.4$, there is a 3.3% increase in dirty production (and thus emissions) and more than a doubling in the relative profitability of clean innovation. Even with a high elasticity of substitution between clean and dirty production ($\epsilon = 10$), we still find a positive effect of removing patent protection on dirty technology on the relative profitability of clean innovation.

In fact, only when the probability of making a successful innovation is unreasonably high (that is, when $\eta_c = \eta_d = 0.4$) do we find a negative effect on the relative profitability of clean innovation from our patent policy.[35] The reason for this is that when scientists have a high chance of making an innovation, a large share of machines will be produced by a monopolist. In that case, removing patent protection on dirty technology will lead to an increase in the market size of dirty production that is sufficiently high to induce innovators to the dirty sector despite the reduced price on dirty innovations.

5.2 Long-run effects

To explore the long-run effects of our patent policy, we remove patent protection on dirty technology in period 1 and then simulate the economy for the full 80 periods, that is, 400 years. As before, we assume that there is no climate policy in place. In general, due to the low initial productivity of clean technology, both innovation and production eventually transition to dirty technology, resulting in a climate catastrophe. The goal of this exercise is to explore whether patent policy can delay (or even avoid) the transition to dirty technology.

Figure 2 shows the simulation for the first 200 years for our baseline parameters and for different assumptions regarding μ. When patent policy is not used ($\mu = 1$) innovation quickly transitions to dirty technology. Eventually, production also transitions to the dirty sector, at which point an environmental disaster occurs. The removal of patent protection on dirty

[35] We consider this to be an unrealistic scenario since a 40% per annum probability of making a successful innovation corresponds to an unreasonably high long-run growth rate of the economy.

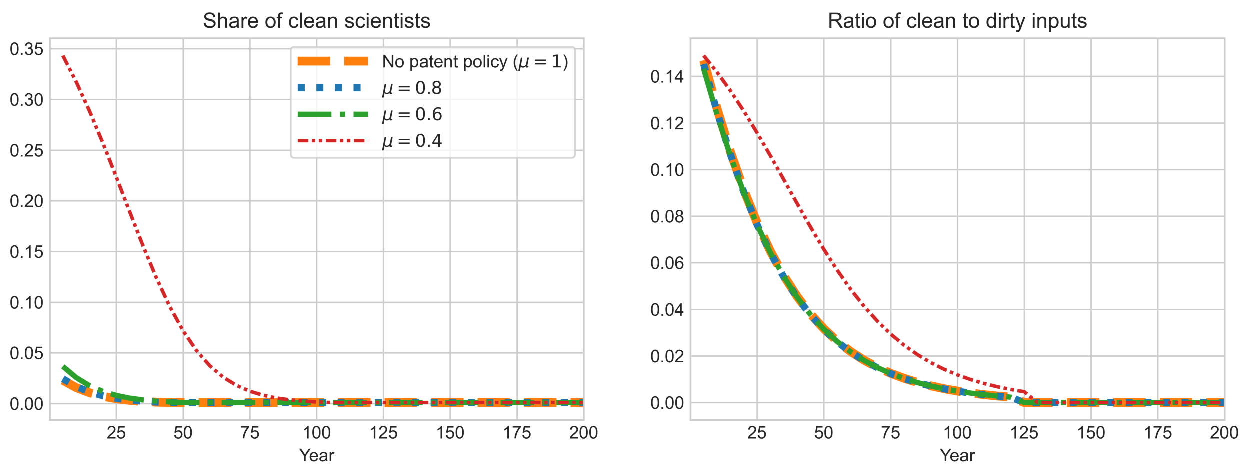

Figure 2: Effect of patent policy on long-run transition to clean technology for different caps on the dirty markup (baseline parameters).

technology increases the relative profitability of clean innovation, which slows the transition of scientists to the dirty sector. For example, when patent policy causes a large reduction in the markup of dirty technology ($\mu = 0.4$), the transition to dirty innovation is delayed by more than 50 years. However, in none of the scenarios is patent policy sufficient to induce the transition to clean technology, which indicates the need to combine patent and climate policies to guarantee the energy transition.

One reason why patent policy leads to only a slight delay in the transition to dirty technology in figure 2 is the large initial technology gap between clean and dirty technology in the baseline simulations. To explore this further, we modify the simulations assuming a + 100 % increase in the initial productivity of clean technology, A_{c0}. Figure 3 shows that although the increase in productivity is not sufficient to induce scientists to the clean sector, it now takes longer for the economy to transition to dirty technology. As expected, using patent policy has a substantially larger impact on the transition than before. In fact, when patent policy leads to a large reduction in dirty markup ($\mu = 0.4$), patent policy becomes sufficient to induce innovation to clean technology and avoid the climate disaster in the long-run.

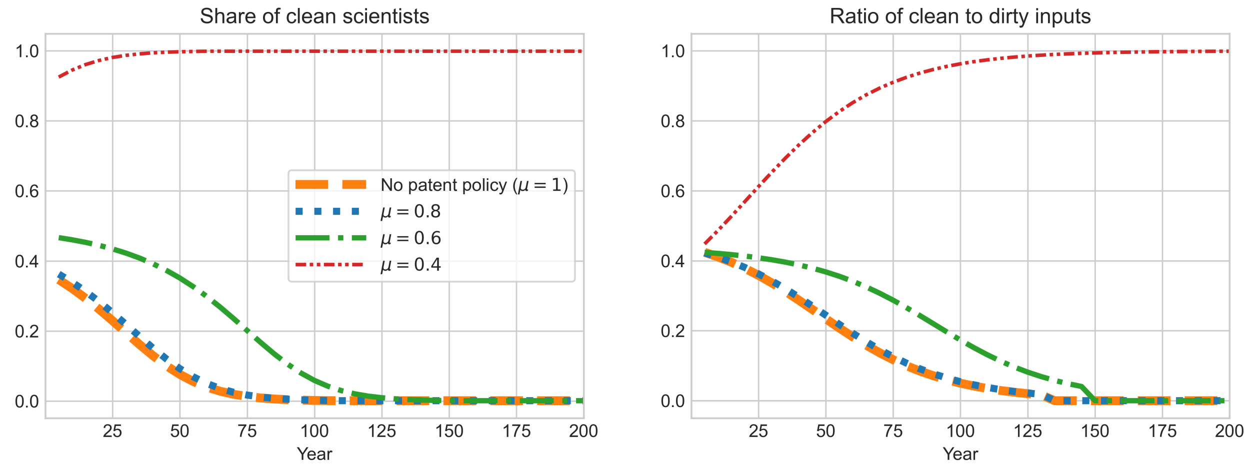

Figure 3: Effect of patent policy on long-run transition to clean technology for different caps on the dirty markup (+100% increase in A_{c0}).

5.3 Implications for climate policy

We now explore the impact of our patent policy on the optimal paths for the carbon tax and the clean innovation subsidy for different assumptions regarding μ. To evaluate the welfare implications, we also find the optimal paths in a scenario with no efficiency losses in climate policy, i.e. $d_q = d_\tau = 0$. We use the level of welfare in this scenario as a benchmark to compare potential welfare gains of using patent policy when there are efficiency losses in the funding of climate policy. As before, we simulate the economy for 80 periods, but show only the first 40 periods since the transition to clean technology is completed by then.

Figure 4 shows the optimal paths for the carbon tax and the clean innovation subsidy for our baseline estimate of the elasticity of substitution, $\epsilon = 3$. In the benchmark scenario (solid gray line) where climate policy can be financed without losses, both the optimal carbon tax and R&D subsidy are steadily increasing over time. However, while the carbon tax remains at a high level, the R&D subsidy is gradually phased out after 50 years. Despite the relatively high carbon tax and R&D subsidy in the benchmark scenario, it still takes approximately 100 years for scientists to have transitioned to the clean sector and twice that time for dirty inputs to be completely substituted by clean inputs. The slow transition results in a considerable increase in the mean temperature.

Introducing efficiency losses in climate policy results in a trade-off between the carbon tax and R&D subsidy. When patent policy is not used, that is, $\mu = 1$ (dashed yellow line), the carbon tax is postponed by more than 50 years, which is compensated by a sharp increase in the R&D subsidy. Once the carbon tax is introduced, the R&D is quickly phased out. Although the delay in carbon taxation initially reduces the share of clean input in production, the high subsidy to clean innovation induces more scientists to the clean sector and speeds up clean productivity growth. However, once the subsidy to clean innovation is phased out, the transition of scientists slows down. The net effect is a delay in the transition to clean innovation.

Removing patent protection on dirty innovations ($\mu < 1$) allows a similar transition to clean technology as before, but for a substantially lower R&D subsidy. In the absence of patent protection on dirty technology, clean innovation becomes relatively more profitable, reducing the need to subsidize clean innovation. In fact, when patent policy results in a dirty markup close to the competitive price, that is, $\mu = 0.4$ (recall that μ is bounded below by $\alpha = 1/3$), the optimal R&D subsidy is even lower than in the benchmark scenario. In general, when environmental policy cannot be financed without losses, there is a delay in the transition to clean innovation. Removing access to patent protection on dirty innovations allows for a reduction in the optimal R&D subsidy, and for large reductions in the dirty markup, there is even a reduction in the optimal carbon tax and a faster transition to clean innovation.

These results are robust to alternative levels of substitutability between clean and dirty inputs. Figures S1 and S2 in Supplementary Material show the optimal transition to clean technology when $\epsilon = 1.5$ (weak substitutes) and $\epsilon = 10$ (very strong substitutes), respectively. When ϵ is low, the carbon tax is delayed slightly longer, but once it is introduced, it must increase steadily over time to ensure that dirty production is eventually phased out. When

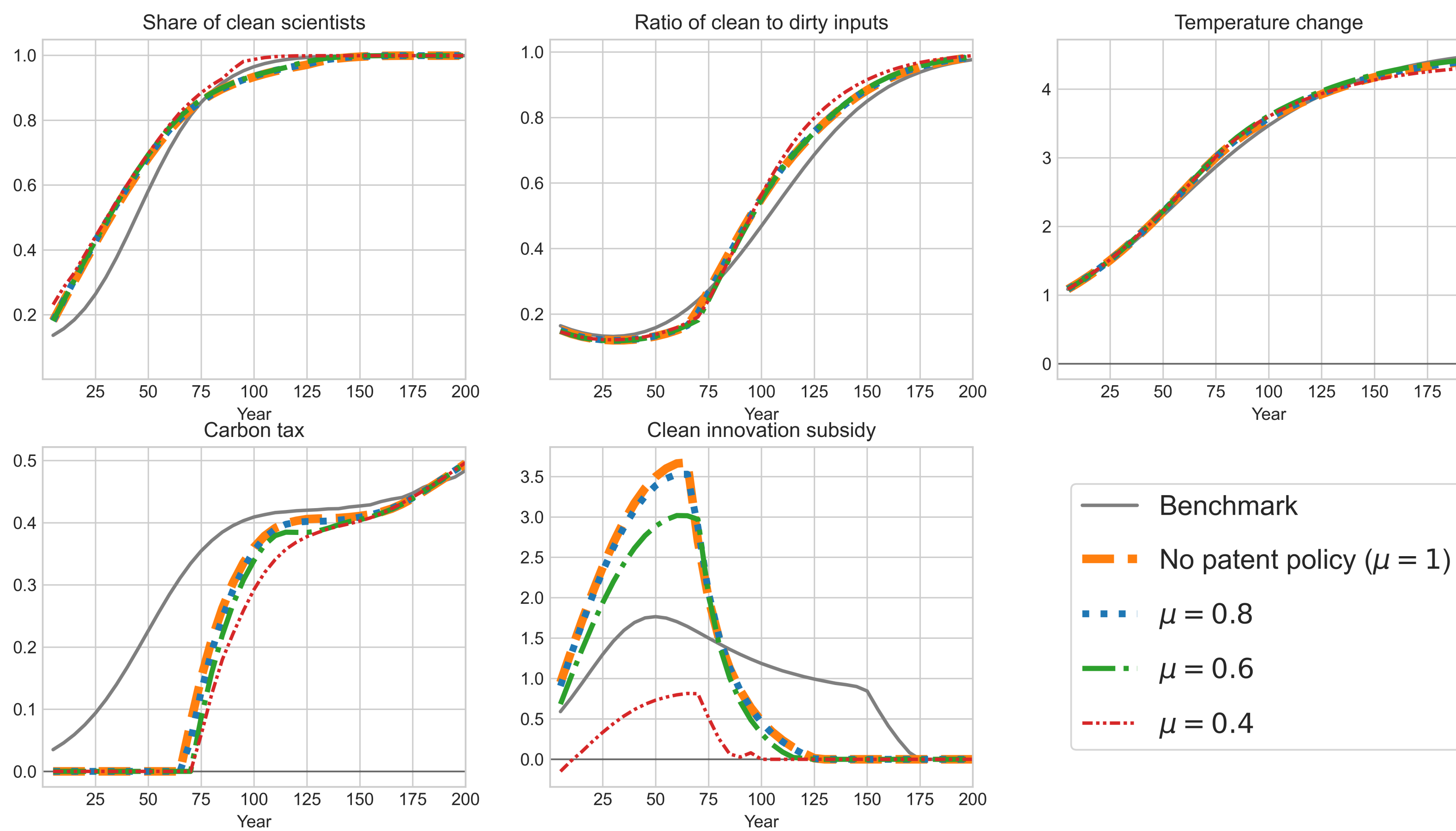

Figure 4: Optimal paths for the carbon tax and clean innovation subsidy and for different caps on the dirty markup (baseline parameters). Benchmark refers to a scenario without efficiency losses in climate policy, $d_\tau = d_q = 0$.

ϵ is high, the transition to clean technology is always faster, and it is induced with only a modest carbon tax combined with a high but short-lived R&D subsidy. However, in all cases, the use of patent policy reduces the required stringency of standard climate policy to ensure the energy transition.

Table 3 shows the welfare gains from combining climate policy with patent policy for different assumptions regarding μ and ϵ. Welfare costs are measured as the equivalent percentage increase in per-period consumption that would be required to make consumers as well off as in the benchmark scenario where climate policy can be financed without losses. Column 1 shows the percentage loss in consumption-equivalent welfare relative to the benchmark for our baseline estimate of ϵ. When patent policy is not used ($\mu = 1$), consumers would require a 5.2% increase in per-period consumption to be equally well off as in the benchmark scenario.[36] Column 2 shows the percentage increase in welfare compared to when patent policy is not used. When $\mu = 0.6$, that is, dirty innovators can charge only 60% of the unrestricted monopoly price in the absence of a patent, the welfare loss is reduced to 4.1%, which translates to a 21.5% recovery of the welfare loss.

Columns 3-6 present the welfare analysis for when $\epsilon = 1.5$ and $\epsilon = 10$. In both cases, our patent policy improves consumer welfare, especially when removing patent protection results in large reductions in the dirty markup. Column 6 shows that even when ϵ is high, there are still significant improvements in welfare caused by our patent policy despite the large increase in the demand for dirty technology in the short-run. However, column 5 shows that the improvement is less pronounced in "absolute" terms since the welfare loss associated with policy distortions is generally lower.

Notice that there are two sources of welfare improvements reported in table 3. First, removing patent protection on dirty technology induces clean innovation, which reduces the necessary carbon tax and R&D subsidy. Second, since our patent policy reduces the markup on dirty technology, there is also a reduction in the monopoly distortion in the dirty sector.

[36] Recall that one period is 5 years in the simulations.

Table 3: Welfare analysis of patent policy for different elasticity of substitution.

	Baseline		Robustness			
Parameter choice:						
Elasticity of substitution (ϵ)	3		1.5		10	
	(1) $\%Cons$	(2) $\%gain$	(3) $\%Cons$	(4) $\%gain$	(5) $\%Cons$	(6) $\%gain$
No patent policy ($\mu = 1$)	5.2	-	5.3	-	3.1	-
$\mu = 0.8$	4.8	6.4	5.1	4.5	3.0	3.0
$\mu = 0.6$	4.1	21.5	4.4	17.0	2.6	15.5
$\mu = 0.4$	2.0	61.6	3.3	37.8	0.9	71.4

Note: $\%Cons$ refers to the percentage loss in consumption-equivalent welfare relative to the benchmark, i.e., when there are no efficiency losses associated with climate policy. $\%gain$ refers to the percentage increase in consumption-equivalent welfare compared to when there is no use of patent policy ($\mu = 1$).

Although our patent policy alleviates some of the monopoly distortion, the main source of welfare recovery in table 3 comes from the reduction in the efficiency losses from standard climate policy (see Section S.3 in Supplementary Material).

6 Extensions

In the policy analysis in the previous section, we assumed that sub-optimal carbon taxes were caused by efficiency losses in the transfer of public funds. However, there are many reasons why carbon prices are too low. One reason could be the strong opposition against carbon taxes among many voters, which makes it difficult for policymakers to increase carbon taxes. Furthermore, global warming was close to 5°C, which is substantially higher than most climate mitigation goals. We therefore explore two extensions of our main model. First, we constrain optimal carbon taxation by assuming an upper limit on the rate of taxation that policymakers are willing to implement. Second, we increase the economic damages of climate change so that optimal global warming in our simulations is closer to the consensus target among climate scientists. We find that our main results regarding patent policy are robust to these extensions.

6.1 Cap on the carbon tax

Instead of an efficiency loss in taxes, we assume that there is a cap on the carbon tax that policymakers are able to implement. As before, we find the path of climate policy that maximizes the discounted sum of utility, but now subject to the constraint that $\tau_t \leq \bar{\tau} \; \forall \; t$. To estimate this cap, $\bar{\tau}$, we assume that policymakers cannot price global emissions at higher levels than what we are currently observing. Sweden has one of the highest carbon taxes in the world, which in 2020 was \$119 per tonne of CO2.[37] However, since only 13% of global emissions are currently taxed, we assume that the carbon tax can not exceed \$15.5 per tonne of CO2 (\$119 $\times$ 0.13).

In our model, the carbon tax is defined as a share of the price of the dirty input. Therefore, we convert the tax to a share of the price of fossil fuels as follows. First, we estimate the price and carbon content of global fossil fuel consumption. Table S4 in Supplementary Material shows the average price and consumption shares of the main fossil fuels (coal, natural gas and oil) in 2016-2020, as well as estimates of the average carbon content for each of the fuels. We estimate the carbon content and price of fossil fuel consumption as weighted averages of fuel-specific carbon contents and prices, respectively, with consumption shares as weights.[38] Second, we adjust the carbon tax by multiplying it with the carbon content of fossil fuel consumption. Third, we divide the adjusted carbon tax with the price of fossil fuel consumption.[39] This results in a cap on the carbon tax equal to 19% of the price of the dirty

[37]The Carbon Pricing Dashboard by the World Bank keeps track of carbon pricing initiatives around the world, and calculates the global coverage rate: https://carbonpricingdashboard.worldbank.org/.

[38]We estimate an average price of fossil fuel consumption of \$6.11 per million Btu, and a carbon content of 74.1 kg CO2 per million Btu of fossil fuels consumed.

[39]More specifically, we estimate the cap on the carbon tax as $\bar{\tau} = \frac{\tau(\$/tCO2)*CC(tCO2/mBtu)}{P_d(\$/mBtu)}$, where τ is the carbon tax in USD per tonne of CO2, CC is the carbon content of fossil fuel consumption in tonne CO2 per million Btu, and P_d is the price of fossil fuel consumption in USD per million Btu.

input in our model.[40]

The simulation results are shown in Section S.4 in Supplementary Material. Figure S3 shows that the carbon tax cap causes the optimal tax to start at a higher rate than in the benchmark scenario, that is, when there are no restrictions on climate policy. The cap quickly becomes binding, and, consequently, the tax level remains equal to the cap throughout the simulation period. To compensate for the low carbon tax, the economy must rely on a higher research subsidy to ensure the transition to clean technology. Removing patent protection on dirty technology ($\mu < 1$) allows a similar transition to clean technology but for lower levels of the R&D subsidy. Table S5 shows the welfare gains associated with patent policy when traditional climate policy is constrained. Once again, we find that the use of patent policy allows for a substantial recovery of the welfare loss associated with climate policy. Therefore, our results from the main analysis are robust to introducing a cap on the carbon tax.

6.2 Higher damages

In the main analysis, due to the initially low productivity of clean technology, dirty technology remained in use for a long time, which caused a substantial increase in the mean temperature. Therefore, we check the robustness of our results to a more stringent climate mitigation goal by tripling the estimate of damages used to calibrate the damage function. We now assume that damages are 30% at 3°C warming.[41] The new calibration of the damage function is consistent with global warming not exceeding 3°C in any of the patent policy scenarios, which is closer to the climate mitigation targets set by countries.

Figure S4 shows that higher damages cause the carbon tax to start at a higher rate compared to the benchmark in figure 4 in the main analysis. An initially higher carbon tax speeds up the transition to clean technology in both production and innovation, which has

[40]For comparison, Acemoglu et al. (2016) estimated a cap of 23% in their model.

[41]This results in $\lambda = 2.7638$. Note that for this value of λ, the damage function becomes almost linear in global warming.

two implications for the future R&D subsidy and carbon tax. First, there is now less need for a clean innovation subsidy than before. Second, the carbon tax does not have to increase to the same level as in the main analysis. However, as in the main analysis, introducing the efficiency losses causes a delay in the carbon tax and an increase in the clean innovation subsidy. Removing patent protection on dirty technology helps foster the transition to clean technology by reducing the necessary carbon tax and R&D subsidy. Table S6 shows that this results in similar welfare gains from patent policy as before.

7 Discussion and conclusion

This paper proposes the novel use of patent policy to foster the transition to clean technology. We contribute to the literature by exploring alternative policy tools to induce clean innovation when policymakers are unable to implement optimal carbon taxation and public funding of research. Patent policy is introduced in an endogenous growth model with environmental constraints to analyze the short- and long-run effects on the energy transition from removing patent protection on dirty technology. Despite the increase in demand for dirty technology in the short-run, we show both analytically and numerically that excluding dirty technology from patent protection can induce clean innovation for a wide range of reasonable parameter assumptions. Although patent policy might not be sufficient to induce the transition to clean technology on its own, our numerical simulations indicate that using patent policy as an additional policy tool can substantially reduce the welfare cost of mitigating climate change when traditional climate policy is constrained.

While global coordination has been a major obstacle for the carbon tax, there has been substantial international collaboration on intellectual property rights since the early 1990s. All members of the World Trade Organization must adhere to the TRIPS Agreement, which establishes a set of minimum standards of protection that member countries must offer. Most countries have national patent offices; however, many countries have come together to form the European Patent Office (EPO), which offers a single patent grant procedure for its members. These patents are known as European patents. In fact, most patent applications are filed at only two patent offices in the world – the USPTO and the EPO. Our suggested policy, namely, to exclude dirty innovations from patent protection, either globally or at local patent offices, implies an unequal treatment of technological innovations. Although this might seem difficult to implement at first glance, in the following lines, we discuss its feasibility and document real-world examples that resemble our suggested policy.

The first challenge to implementing this policy lies in the patent regime. The current regime is based on a uniform system where patent protection is available for inventions in all fields of technology without discrimination. Our proposed policy, on the other hand, requires unequal treatment of innovations based on their environmental impact. Although our policy proposal is novel, the possibility of implementing a differentiated system of patent protection is already being discussed by policy institutions and the law literature (e.g. Gollin, 1991, OECD, 2004, Derclaye, 2008). In fact, there are currently some exceptions to patentability, such as surgical methods and mathematical principles, which have been deemed unethical to patent. This provision in international patent legislation could, in theory, be extended to apply also to technologies that are harmful to the environment.[42]

The second challenge lies in the practical aspect of the implementation. Our policy requires the exclusion of dirty technology from patent protection. Hence, the patent office examiner should not grant a patent to any innovation in a dirty technology. In practice, however, it is likely that innovations are not either clean or dirty, but have attributes in both sectors (or even non-environmental attributes). In such cases, our policy can still be implemented through an evaluation of patent claims. As mentioned previously, patent claims are statements that explain the invention and define which technology is being protected. It

[42]Article 27, paragraph 2 in the TRIPS Agreement states: "Members may exclude from patentability inventions, the prevention within their territory of the commercial exploitation of which is necessary to protect ordre public or morality, including to protect human, animal or plant life or health or to avoid serious prejudice to the environment, provided that such exclusion is not made merely because the exploitation is prohibited by their law."

is common practice that during the patent application process, the patent office examiner rejects some of the claims made by the applicant. This occurs, for instance, when the technology claimed does not fall within the actual invention, that is, the applicant claims a patent that is too broad. Our policy would require an examination by the patent officer to determine the environmental nature of the claims and to reject the dirty ones.[43]

Although our policy has not yet been implemented, there have been some attempts to use patent regulation to promote clean innovation. These have mainly consisted of speeding up the patent examination process for clean innovations by offering these applications a fast-track channel at the patent office. Some examples are the Green Technology Pilot Program that took place at the USPTO between 2009 and 2012, and the Green Channel that was introduced at the UK IPO in 2009. Many national patent offices have implemented similar programs; however, given the low participation rates, it is not clear whether innovators prefer a faster examination process or not (Dechezleprêtre, 2013). Like fast-track programs, our patent policy also requires patent offices to determine the environmental nature of patent applications, but instead of assessing the greenness of the applications, they must now assess their "dirtiness".

References

Acemoglu, Daron, Philippe Aghion, Lint Barrage, and Hemous David. 2019. "Climate Change, Directed Innovation, and Energy Transition: The Long-Run Consequences of the Shale Gas Revolution." *Working paper* .

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. 2012. "The Environment and Directed Technical Change." *American Economic Review* 102 (1):131–166.

[43] A proper examination process should also ensure that innovators are not able to strategically rewrite their patents to make them "look" clean when they are in fact dirty in nature.

Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley, and William Kerr. 2016. "Transition to Clean Technology." *Journal of Political Economy* 124 (1):52–104.

Allen, Myles R. 2016. "Drivers of peak warming in a consumption-maximizing world." *Nature Climate Change 2016 6:7* 6 (7):684–686.

Arora, Ashish, Marco Ceccagnoli, and Wesley M. Cohen. 2008. "R&D and the patent premium." *International Journal of Industrial Organization* 26 (5):1153–1179.

Barrage, Lint. 2014. "Optimal Dynamic Carbon Taxes in a Climate-Economy Model with Distortionary Fiscal Policy." *Working paper* .

Bena, Jan, Miguel A. Ferreira, Pedro Matos, and Pedro Pires. 2017. "Are foreign investors locusts? The long-term effects of foreign institutional ownership." *Journal of Financial Economics* 126 (1):122–146.

Bovenberg, A. Lans and S. Smulders. 1995. "Environmental quality and pollution-augmenting technological change in a two-sector endogenous growth model." *Journal of Public Economics* 57 (3):369–391.

BP. 2022. "BP Statistical Review of World Energy." Tech. rep. http://www.bp.com/statisticalreview.

Bretschger, Lucas, Filippo Lechthaler, Sebastian Rausch, and Lin Zhang. 2017. "Knowledge diffusion, endogenous growth, and the costs of global climate policy." *European Economic Review* 93:47–72.

Chu, Angus C. 2009. "Effects of blocking patents on R&D: A quantitative DGE analysis." *Journal of Economic Growth* 14 (1):55–78.

Dechezleprêtre, Antoine. 2013. "Fast-tracking Green Patent Applications: An Empirical Analysis." Tech. rep., International Centre for Trade and Sustainable Development, Geneva, Switzerland.

Dechezleprêtre, Antoine, Ralf Martin, and Myra Mohnen. 2017. "Knowledge spillovers from clean and dirty technologies." *Grantham Research Institute Working Paper series* 135.

Derclaye, Estaelle. 2008. "Intellectual Property Rights and Global Warming." *Marquette Intellectual Property Law Review* 12 (2):264–297.

Dietz, Simon, Frederick van der Ploeg, Armon Rezai, and Frank Venmans. 2021. "Are Economists Getting Climate Dynamics Right and Does It Matter?" *Journal of the Association of Environmental and Resource Economists* 8 (5):895–921.

Dietz, Simon and Frank Venmans. 2019. "Cumulative carbon emissions and economic policy: In search of general principles." *Journal of Environmental Economics and Management* 96:108–129.

Fischer, Carolyn and Richard G. Newell. 2008. "Environmental and technology policies for climate mitigation." *Journal of Environmental Economics and Management* 55 (2):142–162.

Gallini, Nancy T. 1992. "Patent Policy and Costly Imitation." *The RAND Journal of Economics* 23 (1):52–63.

Gerlagh, Reyer, Snorre Kverndokk, and Knut Einar Rosendahl. 2014. "The optimal time path of clean energy R&D policy when patents have finite lifetime." *Journal of Environmental Economics and Management* 67 (1):2–19.

Gilbert, Richard and Carl Shapiro. 1990. "Optimal Patent Length and Breadth." *The RAND Journal of Economics* 21 (1):106–112.

Gollin, Michael A. 1991. "Using Intellectual Property to Improve Environmental Protection." *Harvard Journal of Law & Technology* 4 (1):193–235.

Greaker, Mads, Tom-Reiel Heggedal, and Knut Einar Rosendahl. 2018. "Environmental Policy and the Direction of Technical Change." *The Scandinavian Journal of Economics* 120 (4):1100–1138.

Greaker, Mads and Lise Lotte Pade. 2009. "Optimal carbon dioxide abatement and technological change: Should emission taxes start high in order to spur R&D?" *Climatic Change* 96 (3):335–355.

Griliches, Zvi. 1990. "Patent Statistics as Economic Indicators: A Survey." *Journal of Economic Literature* 28 (4):1661–1707.

Hall, Bronwyn H and Christian Helmers. 2010. "The role of patent protection in (clean/green) technology transfer." *Santa Clara High Technology Law Journal* 26 (4):487–532.

Hart, Rob. 2019. "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* 6 (2):349–389.

Hegde, Deepak, Alexander Ljungqvist, and Manav Raj. 2021. "Quick or Broad Patents? Evidence from U.S. Startups." *The Review of Financial Studies* 35 (6):2705–2742.

Hémous, David. 2016. "The dynamic impact of unilateral environmental policies." *Journal of International Economics* 103:80–95.

Hori, Takeo and Hiroaki Yamagami. 2018. "Intellectual property rights protection in the presence of exhaustible resources." *Environmental Economics and Policy Studies* 20 (4):759–784.

Jo, Ara. 2020. "Substitution between Clean and Dirty Energy with Biased Technical Change." *Economics Working Paper Series 20/344* .

Kamien, Morton I. and Nancy L. Schwartz. 1974. "Patent Life and R and D Rivalry." *The American Economic Review* 64 (1):183–187.

Karydas, Christos and Lin Zhang. 2019. "Green tax reform, endogenous innovation and the growth dividend." *Journal of Environmental Economics and Management* 97:158–181.

Klemperer, Paul. 1990. "How Broad Should the Scope of Patent Protection Be?" *The RAND Journal of Economics* 21 (1):113–130.

Lanjouw, Jean O., Ariel Pakes, and Jonathan Putnam. 1998. "How to Count Patents and Value Intellectual Property: The Uses of Patent Renewal and Application Data." *Journal of Industrial Economics* 46 (4):405–432.

Lerner, Joshua. 1994. "The Importance of Patent Scope: An Empirical Analysis." *The RAND Journal of Economics* 25 (2):319–333.

Levin, Richard C., Alvin K. Klevorick, Richard R. Nelson, and Sidney G. Winter. 1987. "Appropriating the returns from industrial research and development." *Brookings Paper on Economic Activity* 18 (3):783–831.

Li, Chol-Won. 2001. "On the Policy Implications of Endogenous Technological Progress." *The Economic Journal* 111 (471):C164–C179.

Mansfield, Edwin. 1986. "Patents and Innovation: An Empirical Study." *Management Science* 32 (2):173–181.

Nordhaus, William. D. 1969. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. Cambridge, MA: MIT Press.

O'Donoghue, Ted, Suzanne Scotchmer, and Jacques-François Thisse. 1998. "Patent Breadth, Patent Life, and the Pace of Technological Progress." *Journal of Economics & Management Strategy* 7 (1):1–32.

OECD. 2004. "Patents and Innovation: Trends and Policy Challenges." Tech. rep., OECD Publishing, Paris. https://www.oecd.org/science/inno/24508541.pdf.

Popp, David. 2004. “ENTICE: endogenous technological change in the DICE model of global warming.” *Journal of Environmental Economics and Management* 48 (1):742–768.

Suphaphiphat, Nujin, Pietro F. Peretto, and Simone Valente. 2015. “Endogenous growth and property rights over renewable resources.” *European Economic Review* 76:125–151.

Van Der Wijst, Kaj-Ivar, Francesco Bosello, Shouro Dasgupta, Laurent Drouet, Johannes Emmerling, Andries Hof, Marian Leimbach, Ramiro Parrado, Franziska Piontek, Gabriele Standardi, and Detlef Van Vuuren. 2023. “New damage curves and multimodel analysis suggest lower optimal temperature.” *Nature Climate Change 2023* :1–8.

Zeng, Jinli, Jie Zhang, and Michael Ka Yiu Fung. 2014. “Patent length and price regulation in an R&D growth model with monopolistic competition.” *Macroeconomic Dynamics* 18 (1):1–22.

A Characterization of the decentralized equilibrium

In every period, the final producer has to pay a tax, τ_t, on the price of the dirty input. The maximization problem of the final good producer is

$$\max_{Y_{ct},Y_{dt}} Y_t - p_{ct}Y_{ct} - p_{dt}(1+\tau_t)Y_{dt}$$

where Y_t is given by equation 2. Taking the ratio of the first-order conditions with respect to Y_{ct} and Y_{dt}, we obtain the relative price of the clean input

$$\frac{p_{ct}}{p_{dt}(1+\tau_t)} = \left(\frac{Y_{ct}}{Y_{dt}}\right)^{-\frac{1}{\epsilon}} \tag{20}$$

which is decreasing in the relative supply of the good. Using the price of the final good as the numeraire, the price index of the clean and dirty good is given by

$$p_{ct}^{1-\epsilon} + \left(p_{dt}(1+\tau_t)\right)^{1-\epsilon} = 1 \tag{21}$$

Recall the demand for machines in equation 7 by the intermediate goods producers and the price of machines in equation 10. Since innovators of clean machines are granted a one-period patent on their innovation, they are able to charge the unconstrained monopoly price. Producers of dirty machines, on the other hand, can charge only a share μ of the monopoly markup. Inserting for the price of machines, demand for clean and dirty machines becomes

$$x_{ict}^{MO} = \left(\frac{\alpha^2\Omega(S_t)p_{ct}}{\psi(1-z_c)}\right)^{\frac{1}{1-\alpha}} A_{ict}L_{ct} \quad \text{and} \quad x_{idt}^{MO} = \left(\frac{\alpha^2\Omega(S_t)p_{dt}}{\mu\psi(1-z_d)}\right)^{\frac{1}{1-\alpha}} A_{idt}L_{dt} \tag{22}$$

However, in each period there are only some machines that experience a productivity improvement. The remaining machines are instead sold at the competitive price, i.e. the marginal cost, and there is no need for a production subsidy to correct for the monopoly distortion.

The demand for these machines is given by

$$x_{ijt}^{CO} = \left(\frac{\alpha \Omega(S_t) p_{jt}}{\psi}\right)^{\frac{1}{1-\alpha}} A_{ijt} L_{jt} \tag{23}$$

The number of machines with a successful innovation is given by $n_j s_{jt}^{\sigma}$, while the number of machines sold at the competitive price is given by $1 - n_j s_{jt}^{\sigma}$. Combining these shares with the demand for machines from equations 22 and 23, and using the expression of average machine quality in equation 14, production of the intermediate goods can be written as

$$Y_{ct} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \Omega(S_t)^{\frac{1}{1-\alpha}} p_{ct}^{\frac{\alpha}{1-\alpha}} L_{ct} \widetilde{A_{ct}} \quad \text{and} \quad Y_{dt} = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \Omega(S_t)^{\frac{1}{1-\alpha}} p_{dt}^{\frac{\alpha}{1-\alpha}} L_{dt} \widetilde{A_{dt}} \tag{24}$$

where

$$\widetilde{A_{ct}} \equiv A_{ct} \left[\eta_c s_{ct}^{\sigma} \left(\left(\frac{\alpha}{(1-z_c)}\right)^{\frac{\alpha}{1-\alpha}} - 1\right) + 1\right]$$
$$\text{and} \quad \widetilde{A_{dt}} \equiv A_{dt} \left[\eta_d s_{dt}^{\sigma} \left(\left(\frac{\alpha}{\mu(1-z_d)}\right)^{\frac{\alpha}{1-\alpha}} - 1\right) + 1\right] \tag{25}$$

While A_{jt} is the average quality level of machines in sector j, $\widetilde{A_{jt}}$ is the average effective productivity level that takes into account the fact that some machines in the sector are produced by monopolists. The average effective productivity captures the fact that an increase in innovation efforts will increase not only the average quality of machines, but also the monopoly distortion in the sector. In general, $\widetilde{A_{jt}} < A_{jt}$, with the wedge increasing in the number of scientists in the sector, s_{jt}.[44] Equation 25 shows that a reduction in μ reduces

[44]However, notice that the relative effective productivity of the clean input can still be larger than its relative average quality. For simplicity, assume that there are no machine subsidies ($z_j = 0$) and no patent policy ($\mu = 1$). Then $\frac{\widetilde{A_{ct}}}{\widetilde{A_{dt}}} > \frac{A_{ct}}{A_{dt}}$ if $s_{ct} < \frac{\eta_d^{1/\sigma}}{\eta_c^{1/\sigma} + \eta_d^{1/\sigma}}$. If $\eta_c = \eta_d$, then the relative effective productivity of the clean input is greater than its relative quality level as long as $s_{ct} < 0.5$. In that case, the monopoly distortion is greater in the dirty sector since most innovation efforts are in dirty technology. Therefore, the wedge between $\widetilde{A_{jt}}$ and A_{jt} will be larger in the dirty sector relative to the clean sector.

the wedge between $\widetilde{A_{dt}}$ and A_{dt}. In the absence of patent protection, dirty innovators must charge a lower price for their machines, thus reducing the monopoly distortion in the dirty sector. This leads to the following lemma:

Lemma 1. *All else given, removing patent protection on dirty technology, that is, $\mu < 1$, reduces monopoly distortion in the dirty sector and therefore increases the average effective productivity of dirty technology.*

Proof. $\tilde{A}_{dt}$ is decreasing in μ, as can be seen from its derivative wrt. μ

$$\frac{\partial\left(\left(\eta_d s_{dt}^{\sigma}\left(\left(\frac{\alpha}{\mu}\right)^{\frac{\alpha}{1-\alpha}}-1\right)+1\right)A_{dt}\right)}{\partial\mu} = -A_{dt}\eta_d s_{dt}^{\sigma}\frac{\alpha}{1-\alpha}\left(\frac{\alpha}{\mu}\right)^{\frac{\alpha}{1-\alpha}-1}\frac{\alpha}{\mu^2} < 0$$

□

In addition to machines, the intermediate inputs are produced using also labor. Demand for labor is given by the first-order condition of the intermediate producers problem in equation 6 with respect to L_{jt}

$$(1-\alpha)p_{jt}\Omega(S_t)L_{jt}^{-\alpha}\int_0^1 A_{ijt}^{1-\alpha}x_{ijt}^{\alpha}di - w_t = 0, \tag{26}$$

which results in the following wage rate

$$w_t = (1-\alpha)\Omega(S_t)p_{jt}\frac{Y_{jt}}{L_{jt}} \tag{27}$$

By combining the expressions of intermediate production in equation 24 with the fact that the wage rate in the two sectors must be equal, the relative price of the clean good can be expressed as a function of the effective productivity of clean and dirty technology

$$\frac{p_{ct}}{p_{dt}} = \left(\frac{\widetilde{A_{ct}}}{\widetilde{A_{dt}}}\right)^{-(1-\alpha)} \tag{28}$$

The relative price of the clean input is decreasing in the relative effective productivity of clean technology. Combining this expression of the relative price of the clean input with the price index in equation 21, the price of the clean and dirty inputs can be expressed as functions of the effective productivity and the carbon tax

$$p_{ct} = \frac{\widetilde{A_{dt}}^{1-\alpha}}{\left[\widetilde{A_{ct}}^{\varphi}(1+\tau_t)^{1-\epsilon} + \widetilde{A_{dt}}^{\varphi}\right]^{\frac{1}{1-\epsilon}}} \quad \text{and} \quad p_{dt} = \frac{\widetilde{A_{ct}}^{1-\alpha}}{\left[\widetilde{A_{ct}}^{\varphi}(1+\tau_t)^{1-\epsilon} + \widetilde{A_{dt}}^{\varphi}\right]^{\frac{1}{1-\epsilon}}} \tag{29}$$

where $\varphi = (1-\alpha)(1-\epsilon)$.

Combining the expressions for the relative price of the clean input in equations 20 and 28 together with the expressions for clean and dirty production in equation 24, the relative labor share of the clean input can be expressed as

$$\frac{L_{ct}}{L_{dt}} = \left(\frac{p_{ct}}{p_{dt}(1+\tau_t)}\right)^{-\epsilon} \left(\frac{\widetilde{A_{ct}}}{\widetilde{A_{dt}}}\right)^{-(1-\alpha)} \tag{30}$$

Inserting for the the prices of the clean and dirty inputs from equation 29, the relative labor share can be expressed as a function of the relative effective productivity of the clean input and the carbon tax

$$\frac{L_{ct}}{L_{dt}} = \left(\frac{\widetilde{A_{ct}}}{\widetilde{A_{dt}}}\right)^{-\varphi} (1+\tau_t)^{\epsilon} \tag{31}$$

Since ϵ is positive (and assumed to be larger than one), a higher carbon tax increases the labor share used in clean production. The relative labor share of clean production is also increasing in the relative effective productivity of clean machines. Combining the relative labor share of the clean input with the market clearing condition in the labor market, that is, $L_{ct} + L_{dt} \leq 1$, the labor shares of the clean and dirty inputs can be expressed as functions of effective productivity and the carbon tax

$$L_{ct} = \frac{\widetilde{A_{dt}}^{\varphi}(1+\tau_t)^{\epsilon}}{\widetilde{A_{ct}}^{\varphi} + \widetilde{A_{dt}}^{\varphi}(1+\tau_t)^{\epsilon}} \quad \text{and} \quad L_{dt} = \frac{\widetilde{A_{ct}}^{\varphi}}{\widetilde{A_{ct}}^{\varphi} + \widetilde{A_{dt}}^{\varphi}(1+\tau_t)^{\epsilon}} \tag{32}$$

Recall the expression for profits in the clean sector relative to the dirty sector in equation 18. Inserting for the relative labor share and the relative price of the clean good from equations 31 and 28, we obtain the following expression for the relative profit of clean innovation

$$\frac{\Pi_{ct}}{\Pi_{dt}} = (1+q_t)\frac{\eta_c}{\eta_d}\left(\frac{s_{ct}}{s_{dt}}\right)^{\sigma-1}\frac{1-\alpha}{\mu-\alpha}\left(\frac{\mu(1-z_d)}{1-z_c}\right)^{\frac{1}{1-\alpha}}(1+\tau_t)^{\epsilon}\left(\frac{\widetilde{A_{ct}}}{\widetilde{A_{dt}}}\right)^{-(1+\varphi)}\frac{A_{ct-1}}{A_{dt-1}} \tag{33}$$

where q_t is a research subsidy given to scientists in the clean sector. Inserting for the effective productivity levels and the evolution of technology from equation 15, and assuming that there are no production subsidies ($z_j = 0$), the relative expected profits from clean research is expressed as a function of the share of clean scientists, technology levels and policy instruments.

$$\begin{aligned}\frac{\Pi_{ct}}{\Pi_{dt}} = (1+q_t)\frac{\eta_c}{\eta_d}\left(\frac{s_{ct}}{s_{dt}}\right)^{\sigma-1}\frac{1-\alpha}{\mu-\alpha}\mu^{\frac{1}{1-\alpha}}(1+\tau_t)^{\epsilon}\times \\ \left(\frac{\left(\eta_c s_{ct}^{\sigma}\left(\alpha^{\frac{\alpha}{1-\alpha}}-1\right)+1\right)(1+\gamma\eta_c s_{ct}^{\sigma})}{\left(\eta_d s_{dt}^{\sigma}\left(\left(\frac{\alpha}{\mu}\right)^{\frac{\alpha}{1-\alpha}}-1\right)+1\right)(1+\gamma\eta_d s_{dt}^{\sigma})}\right)^{-\varphi-1}\left(\frac{A_{ct-1}}{A_{dt-1}}\right)^{-\varphi}\end{aligned} \tag{34}$$

The equilibrium allocation of scientists in the decentralized economy is determined by equalizing the expected profits in the two sectors, i.e., $\Pi_{ct} = \Pi_{dt}$.

Clean and dirty production are found by inserting for price and labor from equations 29 and 32 in the expressions for intermediate outputs in equation 24

$$Y_{ct} = \zeta_t \frac{\widetilde{A_{ct}} \widetilde{A_{dt}}^{\alpha+\phi} (1+\tau_t)^{\epsilon}}{\left(\widetilde{A_{ct}}^{\phi}(1+\tau_t)^{1-\epsilon} + \widetilde{A_{dt}}^{\phi}\right)^{\frac{\alpha}{\phi}} \left(\widetilde{A_{ct}}^{\phi} + \widetilde{A_{dt}}^{\phi}(1+\tau)^{\epsilon}\right)}$$

$$\text{and} \quad Y_{dt} = \zeta_t \frac{\widetilde{A_{ct}}^{\alpha+\phi} \widetilde{A_{dt}}}{\left(\widetilde{A_{ct}}^{\phi}(1+\tau_t)^{1-\epsilon} + \widetilde{A_{dt}}^{\phi}\right)^{\frac{\alpha}{\phi}} \left(\widetilde{A_{ct}}^{\phi} + \widetilde{A_{dt}}^{\phi}(1+\tau)^{\epsilon}\right)} \tag{35}$$

where $\zeta_t = \left(\frac{\alpha}{\psi}\right)^{\frac{\alpha}{1-\alpha}} \Omega(S_t)^{\frac{1}{1-\alpha}}$. Inserting for clean and dirty production in equation 2, we can express production of the final good as

$$Y_t = \zeta_t \frac{\widetilde{A_{ct}} \widetilde{A_{dt}} (1+\tau_t)^{\epsilon}}{\left(\widetilde{A_{ct}}^{\phi}(1+\tau_t)^{1-\epsilon} + \widetilde{A_{dt}}^{\phi}\right)^{\frac{\alpha+\epsilon(1-\alpha)}{\phi}} \left(\widetilde{A_{ct}}^{\phi} + \widetilde{A_{dt}}^{\phi}(1+\tau)^{\epsilon}\right)} \tag{36}$$

The total use of machines in sector j is the sum of demand for machines produced by a monopolist and the machines produced competitively

$$\int_0^1 x_{ijt} di = \eta_j s_{jt}^{\sigma} x_{ijt}^{MO} + (1 - \eta_j s_{jt}^{\sigma}) x_{ijt}^{CO}$$

where x_{ijt}^{MO} and x_{ijt}^{CO} are given by equations 22 and 23. Inserting for prices and labor shares from equations 29 and 32, machine use in the clean and dirty sector can be expressed as

$$x_{ct} = \left(\frac{\alpha}{\psi}\right)^{\frac{1}{1-\alpha}} \Omega(S_t)^{\frac{1}{1-\alpha}} \frac{A_{ct}\left[\eta_c s_{ct}^{\sigma}\left(\left(\frac{\alpha}{1-z_c}\right)^{\frac{1}{1-\alpha}} - 1\right) + 1\right] \widetilde{A_{dt}}^{1+\phi} (1+\tau)^{\epsilon}}{\left(\widetilde{A_{ct}}^{\phi}(1+\tau)^{1-\epsilon} + \widetilde{A_{dt}}^{\phi}\right)^{\frac{1}{\phi}} \left(\widetilde{A_{ct}}^{\phi} + \widetilde{A_{dt}}^{\phi}(1+\tau)^{\epsilon}\right)}$$

$$\text{and} \quad x_{dt} = \left(\frac{\alpha}{\psi}\right)^{\frac{1}{1-\alpha}} \Omega(S_t)^{\frac{1}{1-\alpha}} \frac{A_{dt}\left[\eta_d s_{dt}^{\sigma}\left(\left(\frac{\alpha}{\mu(1-z_d)}\right)^{\frac{1}{1-\alpha}} - 1\right) + 1\right] \widetilde{A_{ct}}^{1+\phi}}{\left(\widetilde{A_{ct}}^{\phi}(1+\tau)^{1-\epsilon} + \widetilde{A_{dt}}^{\phi}\right)^{\frac{1}{\phi}} \left(\widetilde{A_{ct}}^{\phi} + \widetilde{A_{dt}}^{\phi}(1+\tau)^{\epsilon}\right)} \tag{37}$$

Aggregating over all the machine lines in the clean sector in equation 11 and inserting for the price and labor share from equations 29 and 32, total profits from the clean sector can be expressed as

$$\Pi_{ct} = (1-\alpha)\left(\frac{\alpha}{1-z_c}\right)^{\frac{1}{1-\alpha}} \zeta_t \frac{A_{ct}\widetilde{A_{dt}}^{\phi+1}(1+\tau)^{\epsilon}}{\left(\widetilde{A_{ct}}^{\phi}(1+\tau_t)^{1-\epsilon}+\widetilde{A_{dt}}^{\phi}\right)^{\frac{1}{\phi}}\left(\widetilde{A_{ct}}^{\phi}+\widetilde{A_{dt}}^{\phi}(1+\tau_t)^{\epsilon}\right)} \tag{38}$$

Both the research subsidy, q_t, and the carbon tax, τ_t, are financed lump sum. However, we assume that there are inefficiencies associated with such government transfers. As a result, only a share $1-d_q$ of the clean research subsidy and $1-d_\tau$ of the carbon tax are rebated to consumers. The market clearing condition for final good thus becomes

$$C_t = Y_t - \psi\left(\int_0^1 x_{ict}di + \int_0^1 x_{idt}di\right) - d_q q_t \Pi_{ct} - d_\tau \tau_t p_{dt} Y_{dt} \tag{39}$$

Since households cannot store the final good, consumption must equal production of the final good net of the amount used up in the production of machines and the amount lost due to inefficiencies in government transfers.

The decentralized equilibrium has now been solved as a function of clean and dirty productivity (A_{ct} and A_{dt}), the carbon tax (τ_t), the clean research subsidy (q_t) and the cap on dirty markup (μ). Assuming that the carbon tax is initially zero and that dirty innovations are granted full protection ($\mu = 1$), using the solutions for clean and dirty production from equation 35, the initial effective productivity levels in the clean and dirty sectors are given by

$$\widetilde{A_{c0}} = \frac{Y_{c0}}{\zeta_0}\left(1+\left(\frac{Y_{c0}}{Y_{d0}}\right)^{\frac{1-\epsilon}{\epsilon}}\right)^{\frac{\alpha+\varphi}{\varphi}} \quad \text{and} \quad \widetilde{A_{d0}} = \frac{Y_{d0}}{\zeta_0}\left(1+\left(\frac{Y_{d0}}{Y_{c0}}\right)^{\frac{1-\epsilon}{\epsilon}}\right)^{\frac{\alpha+\varphi}{\varphi}} \tag{40}$$

and inserting for the effective productivity levels from equation 25, initial average quality of clean and dirty technology can be expressed as

$$A_{c0} = \frac{Y_{c0}}{\zeta_0 \left[\eta_c s_{c0}^{\sigma} \left(\left(\frac{\alpha}{1-z_c} \right)^{\frac{\alpha}{1-\alpha}} - 1 \right) + 1 \right]} \left(1 + \left(\frac{Y_{c0}}{Y_{d0}} \right)^{\frac{1-\epsilon}{\epsilon}} \right)^{\frac{\alpha+\varphi}{\varphi}}$$

$$\text{and} \quad A_{d0} = \frac{Y_{d0}}{\zeta_0 \left[\eta_d (1-s_{c0})^{\sigma} \left(\left(\frac{\alpha}{1-z_d} \right)^{\frac{\alpha}{1-\alpha}} - 1 \right) + 1 \right]} \left(1 + \left(\frac{Y_{d0}}{Y_{c0}} \right)^{\frac{1-\epsilon}{\epsilon}} \right)^{\frac{\alpha+\varphi}{\varphi}} \tag{41}$$

The initial average quality levels, A_{c0} and A_{d0}, are pinned down by the initial production of the two inputs, Y_{c0} and Y_{d0}, and by the initial share of scientists in the clean input sector, s_{c0}.

B Proofs

1. Proof of Proposition 1

Equation 18 in the main text captures the three partial effects on the relative profitability of clean research, where $\frac{p_{ct}}{p_{dt}}$ and $\frac{L_{ct}}{L_{dt}}$ also depend on μ. We proceed with the proof of each partial effect individually:

1. Direct patent policy effect: Given that $\mu \in (\alpha, 1]$, an increase in μ decreases the direct patent policy effect term, $\frac{1-\alpha}{\mu-\alpha}\mu^{\frac{1}{1-\alpha}}$ because

$$\frac{\partial \left[\frac{1-\alpha}{\mu-\alpha}\mu^{\frac{1}{1-\alpha}} \right]}{\partial \mu} = \frac{1-\alpha}{\mu-\alpha}\mu^{\frac{1}{1-\alpha}} \left[-\frac{1}{\mu-\alpha} + \frac{1}{\mu-\mu\alpha} \right] < 0$$

In other words, the direct patent policy effect is increasing in our policy—a lower μ implies higher relative clean profits.

2. Price effect: Combining equations 28 and 25, the price effect term becomes

$$\frac{p_{ct}}{p_{dt}} = \left(\frac{\tilde{A}_{ct}}{\tilde{A}_{dt}}\right)^{-(1\text{-}\alpha)} = \left(\frac{\left(\eta_c s_{ct}^{\sigma}\left(\alpha^{\frac{\alpha}{1-\alpha}}-1\right)+1\right)A_{ct}}{\left(\eta_d s_{dt}^{\sigma}\left(\left(\frac{\alpha}{\mu}\right)^{\frac{\alpha}{1-\alpha}}-1\right)+1\right)A_{dt}}\right)^{-(1-\alpha)}$$

An increase in μ decreases the price effect term because $\tilde{A}_{dt}$ is decreasing in μ (see proof of Lemma 1). In other words, the price effect is increasing in our policy—a lower μ implies higher relative clean profits.

3. Market size effect: From equation 31, the market size effect term becomes

$$\frac{L_{ct}}{L_{dt}} = \left(\frac{\tilde{A}_{ct}}{\tilde{A}_{dt}}\right)^{-(1-\alpha)(1-\epsilon)} (1+\tau_t)^{\epsilon}$$

For any $\epsilon > 1$, an increase in μ increases the market size effect term because $\tilde{A}_{dt}$ is decreasing in μ (see proof of Lemma 1). In other words, the market size effect is decreasing in our policy—a lower μ implies lower relative clean profits.

□

2. Proof of Proposition 2

Equation 34 expresses the relative expected profits of clean research as a function of the share of scientists, technology levels, and the policy instruments. Its derivative with respect to μ determines the effect of our policy on clean relative profits. We assume that there is no climate policy in place ($\tau_t = q_t = 0$). The derivative of equation 34 wrt. μ is given by

$$\frac{\partial\left(\frac{\Pi_{ct}}{\Pi_{dt}}\right)}{\partial\mu}=\frac{\partial\left(\frac{\Pi_{ct}}{\Pi_{dt}}=\frac{\eta_c}{\eta_d}\left(\frac{s_{ct}}{s_{dt}}\right)^{\sigma-1}\frac{1-\alpha}{\mu-\alpha}\mu^{\frac{1}{1-\alpha}}\left(\frac{\left(\eta_c s_{ct}^{\sigma}(\alpha^{\frac{\alpha}{1-\alpha}}-1)+1\right)(1+\eta_c s_{ct}^{\sigma}\gamma)}{\left(\eta_d s_{dt}^{\sigma}((\frac{\alpha}{\mu})^{\frac{\alpha}{1-\alpha}}-1)+1\right)\left(1+\eta_d s_{dt}^{\sigma}\gamma\right)}\right)^{-1\text{-}\varphi}\left(\frac{A_{ct-1}}{A_{dt-1}}\right)^{-\varphi}\right)}{\partial\mu}$$

$$=\frac{\partial\left(\frac{\Pi_{ct}}{\Pi_{dt}}=\kappa\frac{1-\alpha}{\mu-\alpha}\mu^{\frac{1}{1-\alpha}}\left(\frac{\eta_d s_{dt}^{\sigma}((\frac{\alpha}{\mu})^{\frac{\alpha}{1-\alpha}}-1)+1}{\eta_c s_{ct}^{\sigma}(\alpha^{\frac{\alpha}{1-\alpha}}-1)+1}\right)^{1+\varphi}\right)}{\partial\mu}$$

$$=\underbrace{\kappa}_{>0}\underbrace{\left[-\frac{(1-\alpha)}{(\mu-\alpha)^2}\mu^{\frac{1}{1-\alpha}}+\frac{1}{1-\alpha}\mu^{\frac{1}{1-\alpha}-1}\frac{1-\alpha}{\mu-\alpha}\right]}_{<0}\underbrace{\left(\frac{\eta_d s_{dt}^{\sigma}(\left(\frac{\alpha}{\mu}\right)^{\frac{\alpha}{1-\alpha}}-1)+1}{\eta_c s_{ct}^{\sigma}(\alpha^{\frac{\alpha}{1-\alpha}}-1)+1}\right)^{1+\varphi}}_{>0}$$

$$+\underbrace{\kappa}_{>0}\left[(1+\varphi)\underbrace{\left(\frac{\eta_d s_{dt}^{\sigma}(\left(\frac{\alpha}{\mu}\right)^{\frac{\alpha}{1-\alpha}}-1)+1}{\eta_c s_{ct}^{\sigma}(\alpha^{\frac{\alpha}{1-\alpha}}-1)+1}\right)^{\varphi}}_{>0}\underbrace{(-1)\frac{\frac{\alpha}{1-\alpha}\eta_d s_{dt}^{\sigma}\left(\frac{\alpha}{\mu}\right)^{\frac{\alpha}{1-\alpha}-1}\frac{\alpha}{\mu^2}}{\eta_c s_{ct}^{\sigma}(\alpha^{\frac{\alpha}{1-\alpha}}-1)+1}}_{<0}\right]\underbrace{\frac{1-\alpha}{\mu-\alpha}\mu^{\frac{1}{1-\alpha}}}_{>0}$$

where $\kappa\equiv\frac{\eta_c}{\eta_d}\left(\frac{s_{ct}}{s_{dt}}\right)^{\sigma-1}\left(\frac{1+\eta_c s_{ct}^{\sigma}\gamma}{1+\eta_d s_{dt}^{\sigma}\gamma}\right)^{-1\text{-}\varphi}\left(\frac{A_{ct-1}}{A_{dt-1}}\right)^{-\varphi}$. Under $\alpha\in(0,1)$, $\mu\in(\alpha,1]$, $s_{jt}\in[0,1]$ we can determine the signs of almost every element of the derivative, as indicated in the previous expression. The sign of the term $(1+\varphi)$ determines the sign of the second element in the summation, and thus determines the effect of our policy on relative profitability. One can distinguish three cases:

- If $\epsilon=\frac{2-\alpha}{1-\alpha}$, then $\left(1+\underbrace{(1-\alpha)(1-\epsilon)}_{\equiv\varphi}\right)=0$. Hence, the derivative of the relative clean profits wrt. μ is negative, that is, a lower price cap on dirty technology induces scientists to clean innovation.
- If $\epsilon<\frac{2-\alpha}{1-\alpha}$, then $(1+(1-\alpha)(1-\epsilon))>0$. Hence, the derivative of the relative clean profits wrt. μ is negative, that is, a lower price cap on dirty technology induces scientists to clean innovation.

- If $\epsilon > \frac{2-\alpha}{1-\alpha}$, then $(1+(1-\alpha)(1-\epsilon)) < 0$. Therefore, the second element of the derivative is positive, that is, a lower price cap on dirty technology has an ambiguous effect on the relative profitability of clean innovation and thus on scientists.

This proves the proposition.

$\square$

Supplementary material

S.1 Public vs. Private R&D

Data

We use annual country-level data of public and private R&D expenditure from the OECD Science, Technology and R&D Statistics.[45] Due to data limitations on R&D expenditures, our observational level is country-year. As a measure of innovation, we use patent data from PATSTAT.[46] In our benchmark analysis, we use the universe of patent applications registered at the European Patent Office from 1981 to 2016. We focus on those patents produced in OECD countries.[47] Each patent application is assigned to the country of its applicant (or its innovator when there is no applicant). Whenever there are multiple applicants from different countries, we count them fractionally so that each applicant is assigned an equal share of the application.

It is unreasonable to assume that all patents are of equal relevance. To account for the quality of a patent, we weight it by a logarithmic transformation of the number of forward citations, that is, each patent application is multiplied by $log(1 + \#citations)$. Each citation reflects whether the patent of interest is relevant for a later patent, either by the later patent's applicant or by a patent examiner. Each application is assigned to the year of the first filling. Approximately 85% of the patents from 1970 to the present were registered as applications one year after the earliest filling. Therefore, there is a roughly one-year gap between the earliest filing year and the application filing year.[48]

[45] https://www.oecd-ilibrary.org/science-and-technology/data/oecd-science-technology-and-r-d-statistics_strd-data-en

[46] PATSTAT version: Spring 2020.

[47] More concretely, the following countries are included: AU, AT, BE, CL, CA, CN, IS, IL, CZ, DK, EE, FI, FR, DE, GR, HU, IE, IT, JP, KR, LV, LT, LU, MX, NL, NO, PL, PT, TW, NZ, SK, SI, ES, SE, CH, TR, GB, US.

[48] Very few applications have the same year and most of the remaining ones have longer time differences.

Empirical strategy and regression results

In order to compare the effectiveness of each type of R&D expenditure on the creation of new patents, we estimate the following model

$$Pat_{ct} = \beta_0 + \beta_1 PublicR\&D_{ct} + \beta_2 PrivateR\&D_{ct} + X_{ct} + \delta_c + \gamma_t + \epsilon_{ct}$$

where Pat refers to the quality-adjusted number of new patents in country c in year t, either in levels or as log-transformed. $PublicR\&D$ and $PrivateR\&D$ refer to total R&D expenditures funded by the public sector and the private sector, respectively, either in levels or as log-transformed. X includes controls such as GDP, while δ and γ account for country- and time-fixed effects. We wish to test whether $\beta_1 = \beta_2$, which would indicate that public R&D funding has the same effect as private business R&D funding on patenting. We do so using a Wald test.

Table S1 reports our regression results. The first three columns show log-level regressions, while the last three columns show log-log regressions. The p-values of the Wald test are shown in the last row. All specifications control for GDP and include time- and country-fixed effects. Columns 1 and 4 use contemporaneous R&D expenditures, while columns 2 and 5 use a one-year lag. Columns 3 and 5 measure R&D as the average expenditure over the previous three years.

The regressions show a consistent pattern in which R&D expenditures, both of public and private source, are associated with an increase in the number of new patents. More importantly, the point estimates are systematically larger for private expenditures, compared to public expenditures. Significance levels are stronger in the log-log regression. Column 6 shows that a 1% increase in the average public R&D expenditures for the previous three years is associated with a 0.524% increase in patenting. Private R&D expenditure is associate with a larger effect, with a 1% increase in private R&D expenditures associated with a 0.731% increase in patenting. We use the Wald test to determine whether the coefficients are statistically different from each other. The p-values for this test are reported in the last row "WT p-value". Throughout the table, the results for the Wald test do not allow us to reject that the coefficients are different from each other.

Table S1: Relationship between patents (log citation-weighted) and R&D expenditure.

	Log-levels			Log-log		
	(1) Cont	(2) Lag	(3) M3Lag	(4) Cont	(5) Lag	(6) M3Lag
Public R&D	4.23e-6	5.01e-6**	3.71e-6	0.651***	0.628***	0.524***
	(5.8e-6)	(6.08e-6)	(5.8e-6)	(0.120)	(0.116)	(0.110)
Private R&D	1.23e-5***	1.06e-5*	6.79e-6	0.728***	0.630***	0.731***
	(4.47e-6)	(5.05e-6)	(5.24e-6)	(0.161)	(0.153)	(0.159)
Control (GDP)	√	√	√	√	√	√
FE (C)	√	√	√	√	√	√
FE (T)	√	√	√	√	√	√
Observations	934	906	706	934	906	706
Adj. R-squared	0.972	0.973	0.977	0.990	0.989	0.989
Clusters	37	37	35	37	37	35
WT p-value	0.234	0.370	0.566	0.740	0.994	0.291

Note: The dependent variable is the log-transformed number of citation-weighted patents. Clustered standard errors at the country level in parentheses. All regressions are weighted by country average GDP. *** p<0.01, ** p<0.05, * p<0.1.

We choose the model specification in column 6 as our preferred specification. Using the coefficients in column 6, we estimate that there is a 28% efficiency loss in public research funding $(((0.534 - 0.731)/0.731) \times 100)$.

S.2 Calibration of carbon cycle and damage function

The evolution of accumulated emissions, also known as the carbon cycle, is given by equation 5. The increase in CO2 concentrations from one unit of dirty production is given by ξ, which we estimate from the observed value of Y_d and annual CO2 emissions between 2016 and

2020. In this period, global CO2 emissions from energy consumption were 166.77 GtCO2, while global consumption of fossil fuel energy were 2392.11 exajoules. Hence, we set $\xi = 166.77/2392.11$.

An increase in accumulated emissions, S_t, causes an increase in the global mean temperature, $\Delta(T)$, which results in economic damage. The relationship between CO2 emissions and global warming can be approximated by a linear function in accumulated emissions (Dietz and Venmans, 2019, Dietz et al., 2021).[49] We therefore map emissions to temperature change by using the following linear function

$$\Delta(T_t) = \Delta(T_0) + TCRE \times S_t$$

where $\Delta(T_t)$ is the increase in global mean temperature at time t and $\Delta(T_0)$ is the increase in temperatures that we have already experienced at the beginning of our simulations. $TCRE$ is the transient climate response to cumulative carbon emissions, which measures the change in temperatures from an additional unit of accumulated CO2 emission. Dietz et al. (2021) suggests a $TCRE$ of $1.7°C$ per trillion tons of cumulative carbon emissions (TtC). We increase the estimate by 10 % to take into account the warming from non-CO2 greenhouse gases (Allen, 2016). This results in a $TCRE$ of $1.87°C$/TtC, which corresponds to $0.00051°C$/GtCO2. Since our simulations starts after 2020, we assume global warming of $1°C$ at the start of our simulation period.

We follow AABH and relate global warming to economic damages by using the following function

$$\Omega(\Delta(T)) = \frac{(\Delta_{disaster} - \Delta(T))^{\lambda} - \lambda\Delta_{disaster}^{\lambda-1}(\Delta_{disaster} - \Delta(T))}{(1-\lambda)\Delta_{disaster}^{\lambda}},$$

where $\Delta(T)$ is the increase in the global mean temperature above its pre-industrial level given accumulated emissions at time t. We follow AABH and set the disaster temperature, $\Delta_{disaster}$, equal to $6°C$., which corresponds to a maximum carbon budget of $\bar{S}$, of 9,804

[49]See Dietz and Venmans (2019) for an explanation of the underlying geophysics of this relationship

GtCO2. When the carbon budget is used up, damages are a 100% of GDP. In a scenario of $3°C$ temperature increase by 2100, Van Der Wijst et al. (2023) found global damages to be in the range of 10% to 12% of GDP. We therefore match our damage function to correspond to a 10 % reduction in production at $3°C$ warming, which leads to a value of $\lambda = 0.5958$.

Table S2: Parameter values for the carbon cycle and damage function

Parameter	Description	Value
Y_{c0}	Consumption of renewable, nuclear and hydro-power energy in the world from 2016 to 2020 in exajoules, from BP.	464.26
Y_{d0}	Consumption of fossil fuel energy in the world from 2016 to 2020 in exajoules, from BP.	2392.11
$emission_0$	World emissions from energy production from 2016 to 2020 in GtCO2, from BP.	166.77
$TCER$	Transient climate response to cumulative carbon emissions; increase in temperature ($°C$) per unit increase in accumulated emissions (GtCO2).	0.00051
$\Delta(T_0)$	Global warming in Celsius above pre-industrial levels in 2020.	1
$\bar{S}$	Disaster level of accumulated emissions in GtCO2.	$\frac{\Delta_{disaster} - \Delta(T_0)}{TCER} = 9,804$
λ	Parameter to match the AABH damage function to a 10% decrease in production at $3°C$ warming.	0.5958
ξ	Increase in accumulated emissions (GtCO2) from one unit of dirty production.	$\frac{emission_0}{Y_{d0}} = 0.0697$

Note:

S.3 Exploration of the welfare recovery

In the baseline setting, we have two different sources of distortions: one caused by monopoly power in the innovation sector and one caused from the efficiency losses in financing the standard environmental policies, i.e., carbon tax and R&D subsidy. In order to shed some light on the origins of the welfare improvements from patent policy, we conduct an exercise in which we first remove all distortions and then incorporate only one of them. We define a new benchmark as a "first best" scenario in which there is no monopoly distortion or efficiency losses in transfers of public funds. Monopoly distortion is removed by implementing a subsidy for the use of all machines so that the demand for machines is equal to the demand under perfect competition, while the efficiency losses are removed by setting $d_q = d_\tau = 0$. We first compare this benchmark to a scenario in which the monopoly distortion is present, but environmental policy can still be financed without losses. We then introduce our patent policy, assuming different levels of stringency, that is, μ, as before.

Table S3 presents the welfare analysis of this exercise. Once again, our benchmark scenario—the "first best", is such that all distortions are corrected. When the monopoly distortion is not corrected —"No patent policy", the consumption-equivalent welfare loss is equal to 1.4% per period. Note that the welfare loss is considerably lower than the welfare loss in the baseline scenario in Table 3. This suggests that the majority of the loss in the main results is attributable to the cost associated with environmental policy, and not to the monopoly distortion. Consistent with the main results, our policy improves welfare even when there are no efficiency losses associated with environmental policy. This is because of the intrinsic nature of our policy, which removes exclusivity rights to innovations. We take this opportunity to stress the fact that our patent policy, which a priori does not tackle any environmental externality directly, can be used beyond its direct area of influence, i.e., monopoly power, and induce the energy transition.

S.4 More numerical results

Table S3: Welfare analysis of patent policy without efficiency losses in transfers of public funds (and compared to a "First best" scenario without monopoly distortions)

Parameter choice:		
Elasticity of substitution (ϵ)	3	
	(1)	(2)
	$\%Cons$	$\%gain$
No patent policy ($\mu = 1$)	1.4	-
$\mu = 0.8$	1.1	18.7
$\mu = 0.6$	0.8	41.1
$\mu = 0.4$	0.5	63.2

Note: $\%Cons$ refers to the percentage loss in consumption-equivalent welfare relative to the "First best" benchmark, i.e., no monopoly distortion nor efficiency losses in environmental policy. $\%gain$ refers to the percentage increase in consumption-equivalent welfare compared to when there is no use of patent policy ($\mu = 1$). There are no efficiency losses associated with climate policy in any of the patent policy scenarios.

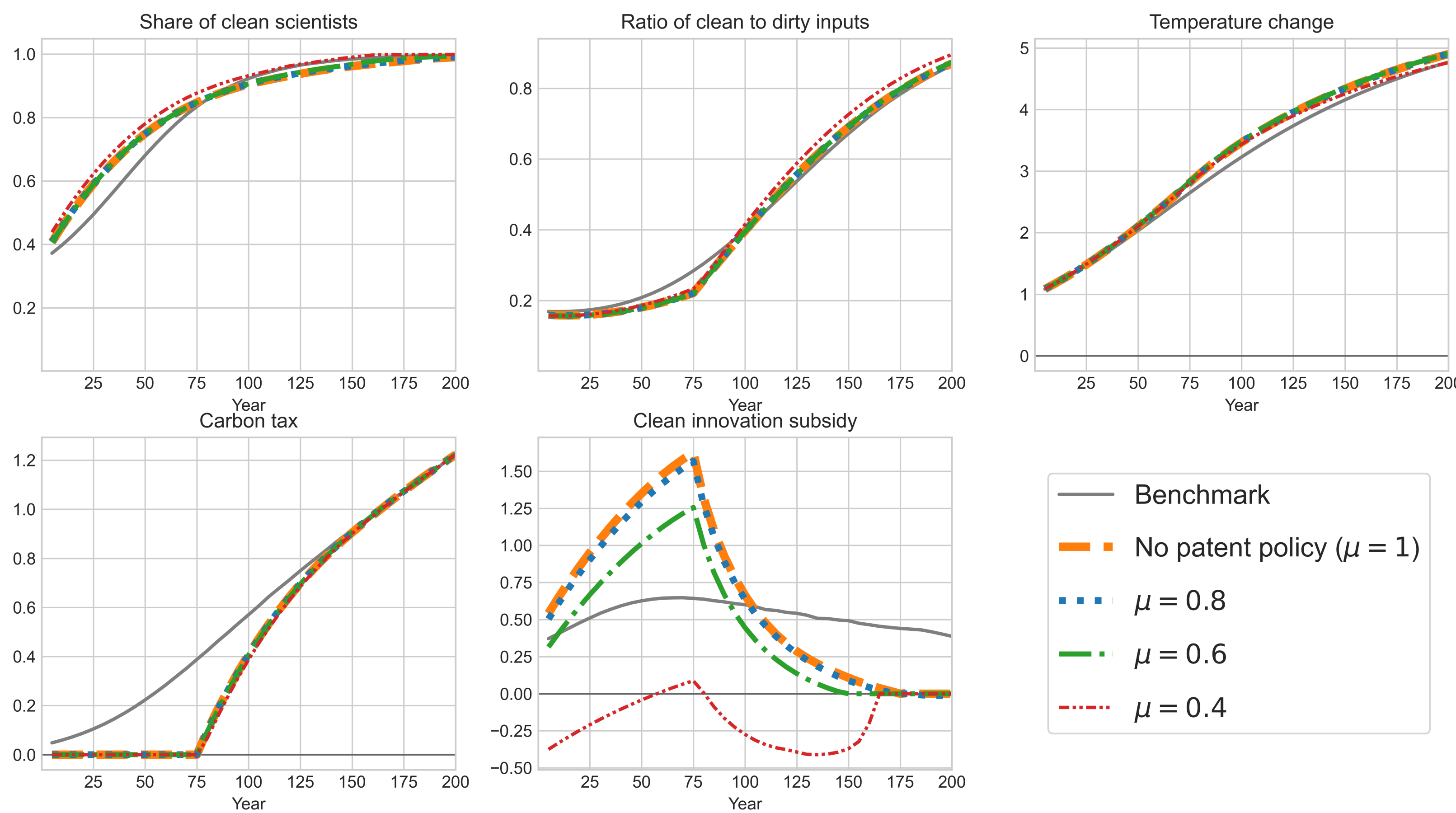

Figure S1: Optimal paths for the carbon tax and clean innovation subsidy and for different caps on the dirty markup (lower $\epsilon = 1.5$). Benchmark refers to a scenario without efficiency losses in climate policy, $d_\tau = d_q = 0$.

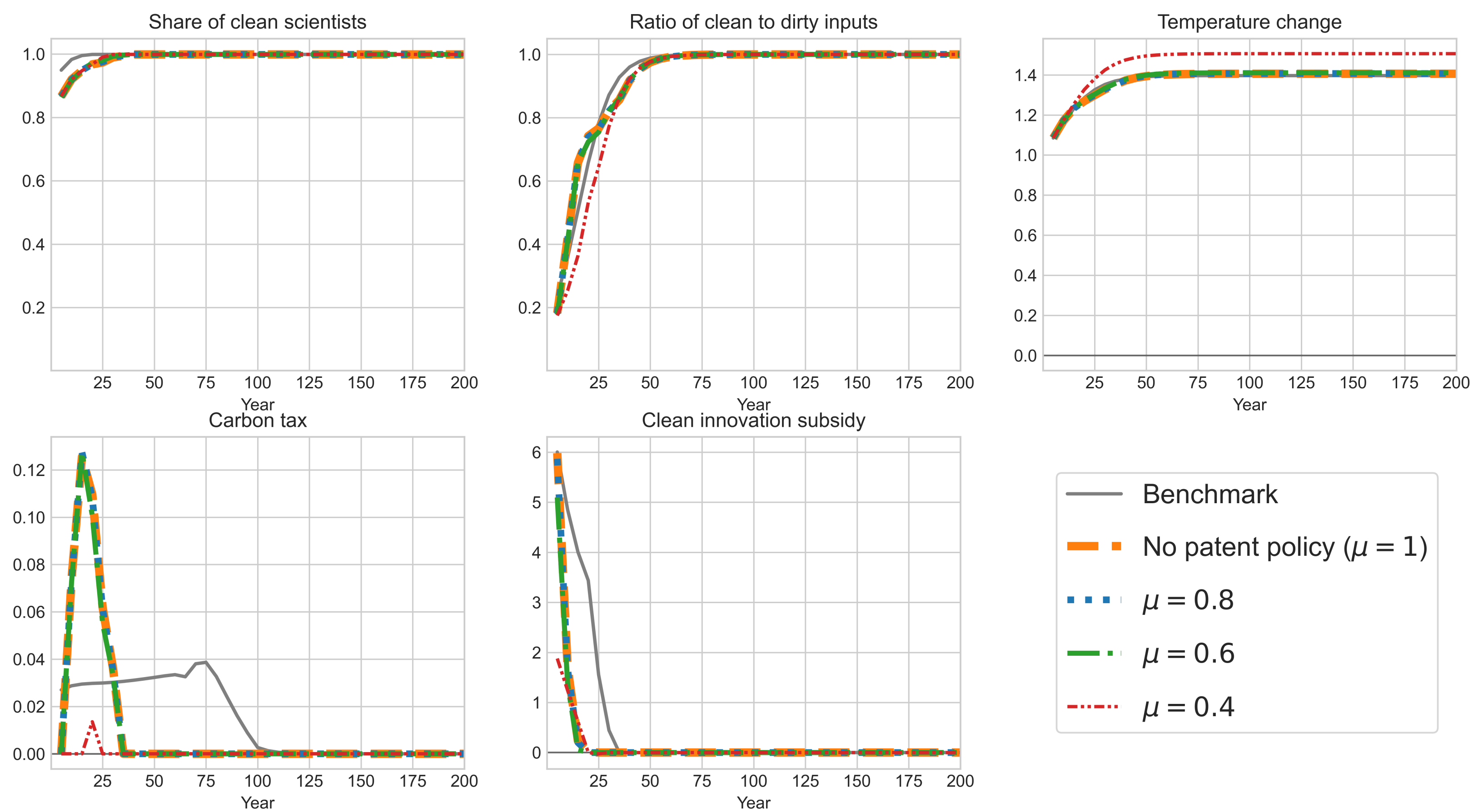

Figure S2: Optimal paths for the carbon tax and clean innovation subsidy and for different caps on the dirty markup (higher $\epsilon = 10$). Benchmark refers to a scenario without efficiency losses in climate policy, $d_\tau = d_q = 0$.

Table S4: Average price, share and carbon content of fossil fuel consumption, 2016-2020.

	(1) Price	(2) Consumption share	(3) Carbon content
Fossil fuel:			
Coal	2.50	0.33	96.10
Natural gas	5.66	0.28	52.91
Oil	9.49	0.39	70.66

Note: Price is measured in USD per million Btu, while carbon content is measured in kg CO2 per million Btu. Price and consumption data is from BP (2022). Average consumption shares are calculated from global consumption of each of the fuels. Average prices are calculated from the Northwest Europe marker price for coal, the average German import price for natural gas, and the Brent crude oil price. Data on carbon content of fossil fuels is taken from: `https://www.eia.gov/environment/emissions/co2_vol_mass.php`.

Table S5: Welfare analysis of patent policy with a cap on the carbon tax.

Parameter choice:		
Elasticity of substitution (ϵ)	3	
	(1) $\%Cons$	(2) $\%gain$
No patent policy ($\mu = 1$)	2.4	-
$\mu = 0.8$	2.1	13.2
$\mu = 0.6$	1.4	42.9
$\mu = 0.4$	-0.3	110.5

Note: $\%Cons$ refers to the percentage loss in consumption-equivalent welfare relative to the benchmark, i.e., when there is no efficiency loss in the R&D subsidy and no cap on the carbon tax. $\%gain$ refers to the percentage increase in consumption-equivalent welfare compared to when there is no use of patent policy ($\mu = 1$).

Table S6: Welfare analysis of patent policy with higher economic damages from global warming.

Parameter choice:		
Elasticity of substitution (ϵ)	3	
	(1) $\%Cons$	(2) $\%gain$
No patent policy ($\mu = 1$)	6.1	-
$\mu = 0.8$	5.8	3.6
$\mu = 0.6$	5.1	15.5
$\mu = 0.4$	2.6	57.0

Note: $\%Cons$ refers to the percentage loss in consumption-equivalent welfare relative to the benchmark, i.e., when there are no efficiency losses associated with climate policy. $\%gain$ refers to the percentage increase in consumption-equivalent welfare compared to when there is no use of patent policy ($\mu = 1$).

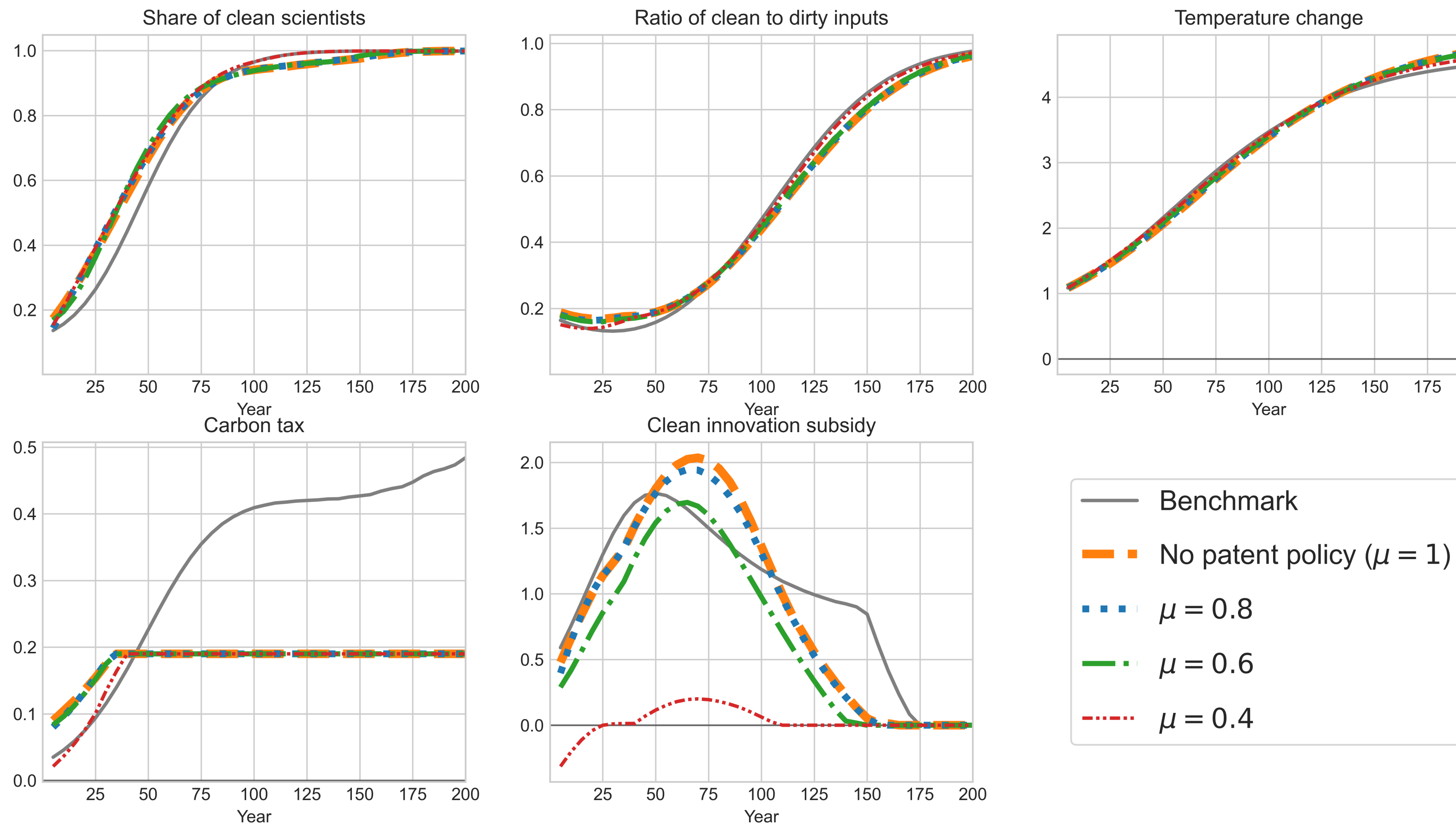

Figure S3: Optimal paths for the carbon tax and clean innovation subsidy and for different caps on the dirty markup (cap on the carbon tax). Benchmark refers to a scenario without efficiency loss in the clean innovation subsidy, $d_q = 0$, and no cap on the carbon tax.

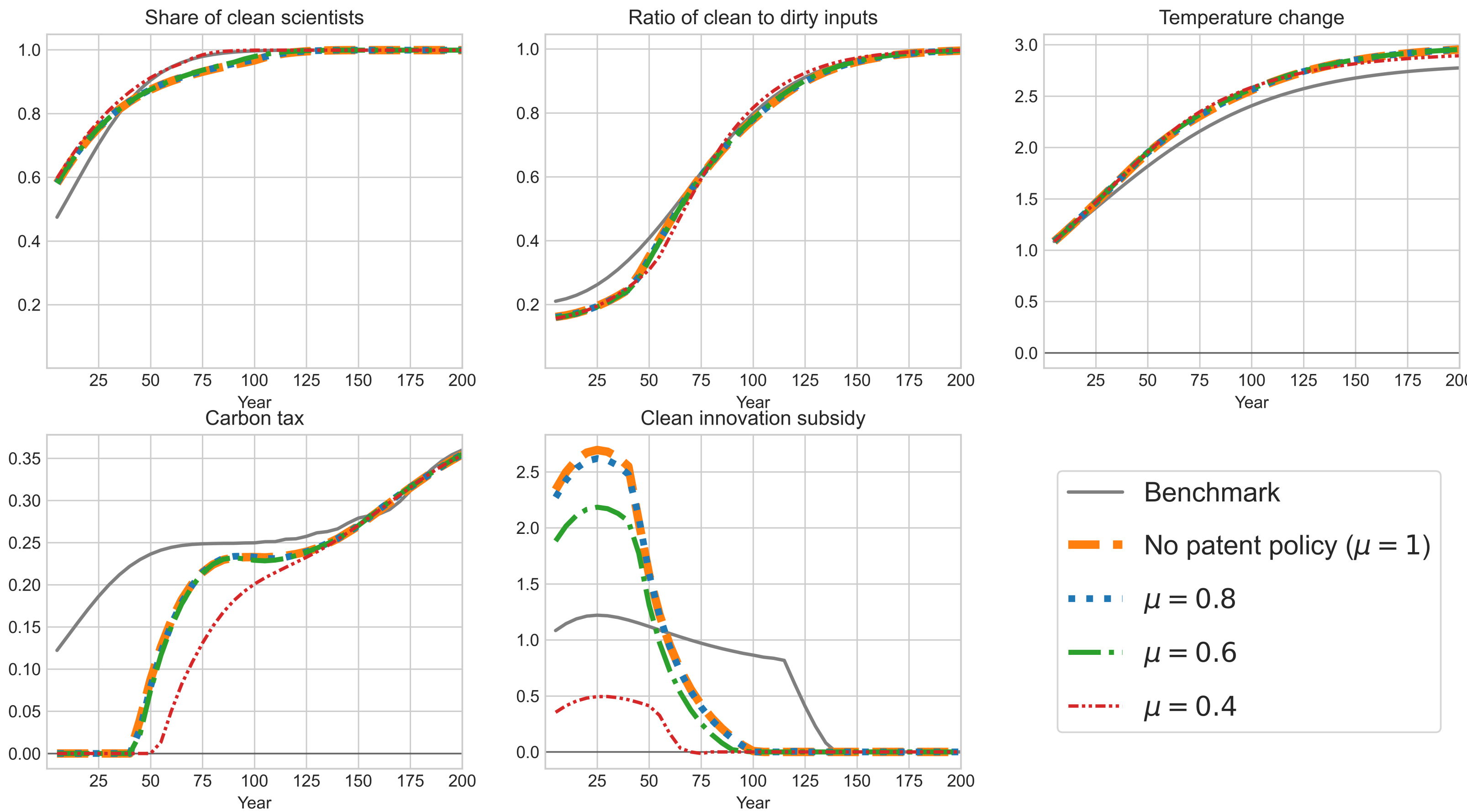

Figure S4: Optimal paths for the carbon tax and clean innovation subsidy and for different caps on the dirty markup (higher damages). Benchmark refers to a scenario without efficiency losses in climate policy, $d_\tau = d_q = 0$.