WPS6859

Policy Research Working Paper 6859

Transition to Clean Capital, Irreversible Investment and Stranded Assets

Julie Rozenberg
Adrien Vogt-Schilb
Stephane Hallegatte

The World Bank
Climate Change Group
Office of the Chief Economist
May 2014

Policy Research Working Paper 6859

Abstract

This paper uses a Ramsey model with two types of capital to analyze the optimal transition to clean capital when polluting investment is irreversible. The cost of climate mitigation decomposes as a technical cost of using clean instead of polluting capital and a transition cost from the irreversibility of pre-existing polluting capital. With a carbon price, the transition cost can be limited by underutilizing polluting capital, at the expense of a loss in the value of polluting assets (stranded assets) and a drop in income. In contrast, policy instruments that focus on redirecting investments—such as feebates or environmental standards—prevent underutilization of existing capital, avoid stranded assets, and reduce short-term losses; but they reduce emissions more slowly and increase the intertemporal cost of the transition. The paper investigates inter- and intra-generational distributional impacts and the political acceptability of climate change mitigation policy instruments.

This paper is a product of the Office of the Chief Economist, Climate Change Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at jrozenberg@worldbank.org.

Transition to Clean Capital, Irreversible Investment and Stranded Assets

Julie Rozenberg [1, 2,*], Adrien Vogt-Schilb [1, 2], Stephane Hallegatte [2]

[1] CIRED, 45bis avenue de la Belle Gabrielle, 94736 Nogent-sur-Marne, France.
[2] The World Bank, Sustainable Development Network, Washington D.C., USA

Keywords: intergenerational equity, energy efficiency standards, mothballing, social and political acceptability, climate mitigation policy
JEL classification: L50, O44, Q52, Q54, Q58

For the past centuries, economic growth has involved the accumulation of fossil-fueled capital that releases greenhouse gases (GHG) in the atmosphere (e.g. coal power plants, gasoline-fueled cars). From a global welfare perspective, this accumulation of polluting capital is sub-optimal because it does not internalize the future economic damages caused by climate change. Stabilizing climate change requires near zero emissions in the long-run, and therefore implies stopping the accumulation of fossil-fueled capital. Future economic growth thus has to rely on clean capital.

In theory, the optimal policy to trigger such a large-scale transition from polluting to clean capital is – in the absence of any other market failure – a Pigouvian price on GHG emissions, for instance a carbon tax (Nordhaus, 1991). However, governments have been timid about the carbon price[1] and have relied instead on instruments that redirect investment towards clean capital, such as energy efficiency standards on new capital (such as the corporate average fuel economy (CAFE) standards in the automobile industry, efficiency standards for new buildings, or feebate programs that tax energy-inefficient equipment and subsidize energy-efficient equipment (IEA, 2014).

In this paper, we investigate how the transition to clean capital is modified when using such *investment-based* instruments instead of a carbon price. We model the accumulation of productive clean capital to replace polluting capital, as suggested by Ploeg and Withagen (1991), and we focus on the effect of the irreversibility of past investments on the transition. The transition to clean capital has been studied before, mostly through the lens directed technical change, focusing on the interaction between pollution and knowledge spillovers externalities (e.g., Gerlagh et al., 2009; Grimaud et al., 2011; Acemoglu et al., 2012).

*Corresponding author
Email addresses: `rozenberg@centre-cired.fr` (Julie Rozenberg), `vogt@centre-cired.fr` (Adrien Vogt-Schilb), `shallegatte@worldbank.org` (Stephane Hallegatte)

[1]Countries that price emissions at a national level are currently limited to the members of the European Union, Switzerland, Kazakhstan, Australia and New-Zealand. Combined with regional or sub-national initiatives, total priced emissions represent less than 10% of global emissions.

Papers by Fischer et al. (2004), Williams (2011), Slechten (2013), and Vogt-Schilb et al. (2014), all study the optimal accumulation of clean capital but do not investigate formally implications for existing polluting capital. Here we discuss the pacing of abatement efforts over time and the impact of different policy instruments on the price of existing capital.

Our analysis builds on a Ramsey model with two types of capital: "polluting" capital, which creates a negative externality (greenhouse gases emissions), and "clean" capital, which does not. Investment is irreversible, meaning that capital can only disappear through depreciation. Firms may however underutilize existing polluting capital, so that abatement efforts can be divided out between two qualitatively different channels: (i) long-term abatement through accumulation of clean capital instead of polluting capital (e.g. people buy electric cars); (ii) immediate abatement through the underutilization (or early-retirement) of polluting capital (e.g. people drive less).

We start from a *laissez-faire* economy, in which marginal productivities of polluting and clean capital are equal. We compare two strategies (carbon price and investment-based instruments) to maintain the concentration of greenhouse gases in the atmosphere below a certain threshold, corresponding to an exogenous policy objective such as the UNFCCC 2°C target.

We find that mitigation costs decompose as a technical cost — using clean instead of polluting capital — and a transition cost due to the irreversibility of pre-existing polluting capital. This irreversibly cost quantifies the regret that society has because of excessive past investment in polluting capital (e.g. having built a coal power plant before the climate mitigation policy has been announced). In the long run both strategies lead to the same steady state, in which most installed capital is clean and GHG concentration is maintained at a constant level. The carbon price and investment-based instruments however induce different trajectories and costs over the short run.

A carbon price maximizes inter-temporal welfare. It redirects all investment towards clean capital until polluting capital has depreciated to a level compatible with the long-term climate ceiling. The carbon price also induces a short-term decrease in the market price of existing polluting capital. This market price has three functions. It values polluting assets, it reflects the demand for new polluting capital and it signals the profitability of polluting capital to investors. If the climate constraint is stringent with regards to past polluting capital accumulation, the market price of polluting capital can even reach zero, and part of this capital is decommissioned. These assets that loose all their value because of the policy are often referred to as "stranded assets" (Goulder et al., 2010; Carbon Tracker, 2013). The underutilization or early retirement of polluting capital allows high short-term abatement but has significant impact on production.

In contrast, investment-based instruments neither create stranded assets nor provide an incentive to underutilize polluting capital. Quite the contrary: by inducing a scarcity effect on polluting capital, these instruments increase the market price of existing polluting capital (i.e. generating "windfall profits" as in Goulder et al., 2010, when emission allowances are grandfathered instead of auctioned). Investment-based instruments yield a higher irreversibility cost than the carbon price as society keeps using obsolete polluting capital until the end of their lifetime instead of early-scrapping it – as if refusing to recognize that past accumulation of polluting capital was a mistake. Thereby, they are less efficient (in inter-temporal welfare terms) than a carbon price.[2] Investment-based instruments lead to a second-best pathway that reaches the same long-run objective as the optimal policy but delays efforts, with lower short-term impacts on output and higher efforts over the medium-run.

Our results highlight a trade-off between the optimality of a climate mitigation policy and its short-term impacts, which may influence political acceptability. If we compare the instruments in terms of welfare maximization, the carbon tax alone is always the best policy. When looking at criteria such as short-term impacts, however, investment-based instruments may appear preferable to some decision-makers and voters. In particular, the impact of the carbon price on asset prices would primarily affect the owners of polluting capital and the workers who depend on them, transforming them into strong opponents to the mitigation policy (Jenkins, In press).

Theoretically, lump-sum cash transfers can compensate the losers and tackle the equity issues faced when implementing a carbon tax (Arrow et al., 1996). In practice, however, it may not be feasible to monitor and compensate each individual loser of climate mitigation policies (e.g., Kanbur, 2010). Another option is to announce a carbon tax in advance to allow economic actors to anticipate it and avoid stranded assets (Williams, 2011), but doing so is made difficult by the governments' limited ability to commit (Kydland and Prescott, 1977; Helm et al., 2003). Finally, one can use a cap-and-trade system where free emission allowances are distributed based on past emissions (grandfathering) or production capacity (Goulder et al., 2010). In this paper, we rethink investment-based instruments as a way to avoid stranded assets, therefore easing the political economy of climate mitigation.

By spreading the costs over time and economic agents, investment-based instruments may reduce the number of opponents to mitigation policies and make the implementation of the carbon tax easier in the long-run. They however cannot curb emissions as fast as the carbon price can. If governments are not able or willing to implement a carbon tax in the near future and the transition has to be triggered by investment-based instruments for political reasons, their slowness makes their implementation (and enforcement) all the more urgent.

The remainder of the paper is structured as follows. Section 1 presents the model and section 2 solves for the *laissez-faire* equilibrium. In section 3 we analyze the optimal growth path, that can be obtained with a carbon price, and we compare it with investment-based second-best instruments in section 4. In section 5, we study the timing issues of investment-based instruments and risks of lock-in. Section 6 discusses the results and concludes.

1. Model

We consider a Ramsey framework with a representative infinitely-lived household, who receives the economy's production from firms y_t, saves by accumu-

[2] A large literature explores their drawbacks — such as the rebound effect if lower energy intensity leads to more extensive use of equipment (Goulder and Parry, 2008; Anderson et al., 2011) — and their rationale such as Tsvetanov and Segerson (2013). Parry et al. (In press) and Allcott (2013) find however that estimated mis-perceptions of energy savings are too low to justify CAFE standards in the automobile sector.

lating assets[3], receives income on assets at interest r_t and purchases goods for consumption c_t. Their wealth thus evolves as:

$$\dot{a}_t = r_t \cdot a_t + y_t - c_t \tag{1}$$

At time t, consuming c_t provides consumers with a utility $u(c_t)$. The utility function is increasing with consumption, and strictly concave ($u' > 0$ and $u'' < 0$).
The household maximizes intertemporal discounted utility W, given by:

$$W = \int_0^\infty e^{-\rho t} \cdot u(c_t)\, dt \tag{2}$$

where ρ is the rate of time preference.

Firms produce one final good y_t, using two types of available capital: polluting capital k_p (e.g., coal power plants, thermal engine vehicles) and clean capital k_c (renewable or nuclear power, electric vehicles).[4]

Production is used for consumption (c_t) and investment ($i_{p,t}$ and $i_{c,t}$).

$$y_t = c_t + i_{p,t} + i_{c,t} \tag{3}$$

Investment $i_{p,t}$ and $i_{c,t}$ increase the stock of installed capital, which depreciates exponentially at rate δ:[5]

$$\dot{k}_{p,t} = i_{p,t} - \delta\, k_{p,t} \tag{4}$$

$$\dot{k}_{c,t} = i_{c,t} - \delta\, k_{c,t} \tag{5}$$

The doted variables represent temporal derivatives.

Investment is irreversible (Arrow and Kurz, 1970):[6]

$$i_{p,t} \geq 0 \tag{6}$$

$$i_{c,t} \geq 0 \tag{7}$$

This means that for instance, a coal plant cannot be turned into a wind turbine, and only disappears through depreciation. However, firms may use only a portion q_t of installed capital k_t to produce the flow of output y_t given by:

$$y_t = F(A_t, q_{p,t}, q_{c,t}) \tag{8}$$

$$q_{p,t} \leq k_{p,t} \tag{9}$$

$$q_{c,t} \leq k_{c,t} \tag{10}$$

F is a classical production function, with decreasing marginal productivities, to which we add the assumption that capital can be underutilized. A_t is exogenous technical progress, and increases at an exponential rate over time.

[3] Assets are capital and loans to other households.

[4] k_p and k_c may also be interpreted as intangible capital; for instance, clean capital encompasses existing clean technologies (e.g. clean electricity production and electrification of the economy) as well as patents, research and development expenses and human capital necessary to develop new clean technologies.

[5] We used the same depreciation rate for polluting and clean capital to keep notations simple, but this assumption plays no particular role in the analysis.

[6] Following the wording by Arltesou (1999) and Wei (2003) capital is putty-clay.

In the remaining of this paper, q_t will be called utilized capital and k_t installed capital. Although it is never optimal in the *laissez-faire* equilibrium, the underutilization of installed polluting capital can be optimal when a carbon price is implemented.[7] For instance, coal plants can be operated part-time and low-efficiency cars can be driven less if their utilization is conflicting with the climate objective.

Polluting capital used a time t emits greenhouse gases $(G \times q_{p,t})$ which accumulate in the atmosphere in a stock m_t.[8] GHG atmospheric concentration increases with emissions, and decreases at a dissipation rate ε:[9]

$$\dot{m}_t = G \cdot q_{p,t} - \varepsilon m_t \tag{11}$$

Note that since emissions are a function of polluting capital and capital has a decreasing marginal productivity, the carbon intensity of output increases with the polluting capital stock.

2. *Laissez-faire* equilibrium

In the *laissez-faire* equilibrium, intertemporal utility maximization leads to a classical arbitrage equation which gives the basic condition for choosing consumption over time (Appendix A):

$$\frac{\dot{c}}{c} = \frac{-u'(c)}{c \cdot u''(c)} \cdot (r_t - \rho) \tag{12}$$

As the elasticity of substitution is positive ($\frac{-u'(c)}{cu''(c)} > 0$), consumption grows if the rate of return to saving r_t is higher than the rate of time preference ρ.

Firms rent the services of polluting and clean capital from households at respective rental rates $R_{p,t}$ and $R_{c,t}$. The flow of profit is given by:

$$\Pi_t = F(A_t, q_{p,t}, q_{c,t}) - R_{c,t} \cdot k_{c,t} - R_{p,t} \cdot k_{p,t} \tag{13}$$

A competitive firm takes $R_{c,t}$ and $R_{p,t}$ as given and maximizes its profit by using all installed capital, equalizing at each time t the marginal productivity of polluting and clean capital to their respective rental rates:

$$\partial_{q_b} F(q_{p,t}, q_{c,t}) = R_{p,t}$$
$$\partial_{q_g} F(q_{p,t}, q_{c,t}) = R_{c,t}$$

The classical equilibrium in capital markets in a Ramsey model applies:

Proposition 1. *In the laissez-faire equilibrium, households are indifferent between investing in polluting or clean capital or lending to other households, so that the marginal productivities of clean and polluting capital net of depreciation are both equal to the interest rate :*

$$R_{p,t} = R_{c,t} = r_t + \delta \tag{14}$$

[7]In this paper, underutilization of clean capital is never optimal: $\forall t,\ q_{c,t} = k_{c,t}$.

[8]In the remaining of the paper, "carbon" will refer to GHG.

[9]The dissipation rate allows maintaining a small stock of polluting capital in the steady state. The linear relation between polluting capital and pollution emission is not a necessary assumption but simplifies the notations.

In the next section, we find that the carbon price forces the marginal productivity of polluting capital to be higher than that of clean capital. Also, because investment is irreversible, the relative price of polluting capital decreases during the transition. We then discuss implications for the political economy of climate mitigation policies.

3. Discounted welfare maximization: carbon price

In this section, we adopt a cost-effectiveness approach (Ambrosi et al., 2003) and analyze policies that allow maintaining atmospheric concentration m_t below a given ceiling $\bar{m}$:

$$m_t \leq \bar{m} \tag{15}$$

This threshold can be interpreted as a tipping point beyond which the environment (and output) can be highly damaged, or as an exogenous policy objective such as the UNFCCC "2°C target" (Allen et al., 2009; Matthews et al., 2009).

We solve for the welfare maximization program, in which institutions internalize the GHG ceiling constraint. A social planner maximizes intertemporal utility given the constraints set by the economy budget, the capital motion law, investment irreversibility and the GHG ceiling. The same strategy can be decentralized by imposing the shadow price of emissions on producers and consumers through an optimal carbon tax or a universal cap-and-trade system (Appendix C).

The social planner program is:

$$\max_{c,i,k} \int_0^\infty e^{-\rho t} \cdot u(c_t)\, dt \tag{16}$$

$$\begin{aligned}
\text{subject to } & F(q_p, k_c) - c_t - i_{p,t} - i_{c,t} = 0 && (\lambda_t)\\
& \dot{k}_{p,t} = i_{p,t} - \delta k_{p,t} && (\nu_t)\\
& \dot{k}_{c,t} = i_{c,t} - \delta k_{c,t} && (\chi_t)\\
& \dot{m}_t = G\, q_{p,t} - \varepsilon m_t && (\mu_t)\\
& m_t \leq \bar{m} && (\phi_t)\\
& i_{p,t} \geq 0 && (\psi_t)\\
& q_{p,t} \leq k_{p,t} && (\beta_t)
\end{aligned}$$

We indicated in parentheses the co-state variables and Lagrangian multipliers (chosen such that they are positive): λ_t is the value of income, ν_t and χ_t are the prices of polluting and clean capital, and μ_t is the price of carbon, expressed in terms of utility at time t. The present value Hamiltonian associated to the maximization of social welfare can be found in Appendix B.

The main first-order conditions of our problem are (Appendix B.1):

$$u'(c_t) = \lambda_t = \nu_t + \psi_t = \chi_t \tag{17}$$

$$\partial_{k_c} F = \frac{1}{\lambda}\left((\delta + \rho)\chi_t - \dot{\chi}_t\right) \tag{18}$$

$$\beta_t = \frac{1}{\lambda}\left((\delta + \rho)\nu_t - \dot{\nu}_t\right) \tag{19}$$

$$\partial_{q_p} F = \beta_t + \tau_t \cdot G \tag{20}$$

Where τ is the price of carbon expressed in dollars per ton:

$$\tau_t = \frac{\mu_t}{\lambda_t} \tag{21}$$

Before the ceiling on atmospheric GHG is reached, a classical result (e.g., Goulder and Mathai, 2000, footnote 11) is that the carbon price exponentially grows at the endogenous interest rate r_t plus the dissipation rate of GHG (Appendix B.3):

$$\forall t,\ m_t < \bar{m} \implies \dot{\tau}_t = \tau_t\,(r_t + \varepsilon) \tag{22}$$

The steady state is reached when $m_t = \bar{m}$. In the steady state, atmospheric emissions are stable, implying that polluting capital is constant at $k_{p,t} = \bar{m}\,\varepsilon/G$ ($\dot{m}_t = 0$, eq. 11) and the rest of the economy keeps growing on a balanced growth path, thanks to exogenous technical change A_t.

In equations 18 and 19 we recognize the rental rates of clean and polluting capital $R_{c,t}$ and $R_{p,t}$, as defined by Jorgenson (1967):

$$R_{c,t} := \frac{1}{\lambda}[(\delta + \rho)\chi_t - \dot{\chi}_t] \tag{23}$$

$$R_{p,t} := \frac{1}{\lambda}[(\delta + \rho)\nu_t - \dot{\nu}_t] \tag{24}$$

where χ_t and ν_t are respectively the clean and polluting capital shadow prices.

The following proposition can be deduced from the first-order conditions:

Proposition 2. *Along the optimal path, the marginal productivity of clean capital equals the rental rate of clean capital:*

$$\partial_{k_c} F = R_{c,t} \tag{25}$$

The marginal productivity of polluting capital is equal to the rental rate of polluting capital plus the marginal cost of carbon emissions:

$$\partial_{q_p} F = R_{p,t} + \tau_t\, G \tag{26}$$

Proof. Equation 25 derives from eq. 18 and 23. Equation 26 is obtained by substituting β_t in eq. 20, using eq. 24. □

In the *laissez-faire* equilibrium, the marginal productivity of polluting capital was also equal to its rental rate. This is no longer the case when the pollution externality is internalized, as firms have to pay the carbon tax when they use polluting capital. Also, the rental rate of polluting capital $R_{p,t}$ is no longer equal to that of clean capital, as it is now affected by an irreversibility cost:

Proposition 3. *Along the optimal path, the interest rate r_t that arbitrates between consumption and investment is given by:*

$$r_t = R_{c,t} - \delta \tag{27}$$

The rental rate of polluting capital can be lower than that of clean capital:

$$R_{p,t} = R_{c,t} - p_t \tag{28}$$

Where the irreversibility cost p_t is the monetary impact of the irreversibility constraint on the rental rate of polluting capital:

$$p_t = \frac{1}{\lambda_t}\left((\rho + \delta)\psi_t - \dot{\psi}_t\right) \in [0,\ R_{c,t}] \tag{29}$$

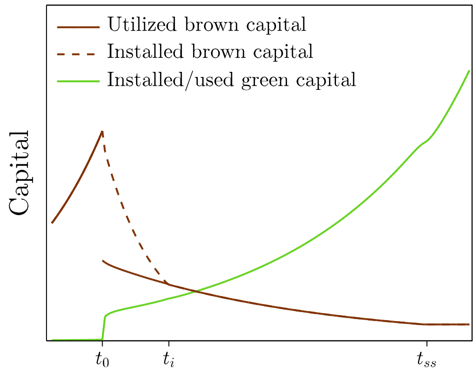

Figure 1: Polluting and clean installed capital, and utilized polluting capital in the first-best optimum. Before t_0, the economy is on the *laissez-faire* equilibrium. At t_0 the carbon price is implemented and polluting capital depreciates until t_i ($\forall t \in (t_0, t_i),\ i_b = 0$). During this period, polluting capital may be underutilized ($q_{p,t} < k_{p,t}$). Polluting investment then starts again, and the steady state is reached at t_{ss}.

Proof. See Appendix B.2 for eq. 27. Equation eq. 28 is obtained by replacing ν_t by $\chi_t - \psi_t$ (eq. 17) in eq. 24. Since $R_{p,t} = \beta_t \geq 0$ (eq. 19), $p_t = R_{c,t} - R_{p,t} \leq R_{c,t}$. □

Because investment is irreversible, when the policy is implemented the stock of polluting capital cannot be instantaneously adjusted to its optimal level. Polluting capital therefore becomes overabundant and its rental rate decreases.

The irreversibly cost p_t quantifies the regret that society has because of excessive past investment in polluting capital (e.g. having built a coal power plant before the climate mitigation policy has been announced). It allows decomposing the shadow price of emissions τ_t as a "technical" abatement cost (e.g. renewable power plants are more expensive than coal power plants) plus an irreversibility cost:[10]

$$\underbrace{\tau_t}_{\text{economic cost}} = \underbrace{\frac{\partial_{q_p} F - \partial_{k_c} F}{G}}_{\text{technical cost}} + \underbrace{\frac{p}{G}}_{\text{irreversibility cost}} \tag{30}$$
$$\text{with } p \in [0,\ \partial_{k_c} F]$$

The next proposition shows that an irreversibility cost necessarily appears when the carbon tax is implemented (in t_0).[11] Once polluting capital has adjusted through natural depreciation, the irreversibility cost is nill.

Proposition 4. *Two phases can be distinguished during the optimal transition to clean capital:*

[10] Combining equations 25, 26 and 28.

[11] Contrary to (Arrow and Kurz, 1970), who find that the irreversibility constraint is binding if the initial capital is higher than the steady-state level, here the irreversibility constraint is binding for any level of initial polluting capital because of the new constraint on emissions.

- *First, a phase when the market price of polluting capital is lower than that of clean capital and no investment is made in polluting capital:*

$$\begin{aligned}
&0 < p_t \leq R_{c,t} \\
&R_{p,t} < R_{c,t} \\
&i_{p,t} = 0
\end{aligned}$$

- *Then, a phase when the market price of polluting and clean capital are equal and polluting investment is strictly positive:*

$$\begin{aligned}
p_t &= 0 \\
R_{p,t} &= R_{c,t} \\
i_{p,t} &> 0
\end{aligned}$$

  Note that during this phase net investment is negative (when accounting for depreciation) until it is equal to zero in the steady-state state.

PROOF. Appendix B.4. □

During the first phase, the irreversibility constraint prevents the economy from transforming polluting capital into either clean capital or consumption and the market price of polluting capital falls below the marginal utility of consumption (eq. 17).

The maximum value of the irreversibility cost p_t is $\partial_{k_c} F$, the marginal productivity of clean capital, as the maximum regret cost is the cost of not having invested in clean instead of polluting capital before t_0. Indeed, if $p_t = \partial_{k_c} F$, the rental rate of polluting capital falls down to zero (eq. 28), reflecting that polluting capital is overabundant:

Proposition 5. *During the first phase (when polluting investment is nill) if the carbon price is higher than the marginal productivity of installed polluting capital, polluting capital is underutilized until its marginal productivity equals the carbon price:*

$$\tau_t \, G > \partial_{k_p} F(k_p, k_c) \implies \begin{cases} p_t = R_{c,t} \\ R_{p,t} = 0 \\ q_{p,t} < k_{p,t} \\ \partial_{q_p} F(q_p, k_c) = \tau_t \, G \end{cases} \tag{31}$$

PROOF. Eq. 26 implies that the rental rate of polluting capital $R_{p,t}$ is the difference between the marginal productivity of polluting capital and the carbon price. As the rental rate of polluting capital $R_{p,t}$ is equal to the positive multiplier associated to the capacity constraint β_t (eq. 19 and 24), when the carbon price is higher than the marginal productivity of installed polluting capital the rental rate of polluting capital is nill and capital is underutilized. □

The underutilization of brown capital depends on the ceiling $\bar{m}$, on the initial stock of brown capital k_{b,t_0} and on other parameters of the model such as the functional forms of F and u, on the depreciation rate δ and the preference for the present ρ. As illustrated in Fig. 2, for a given set of functions and parameters

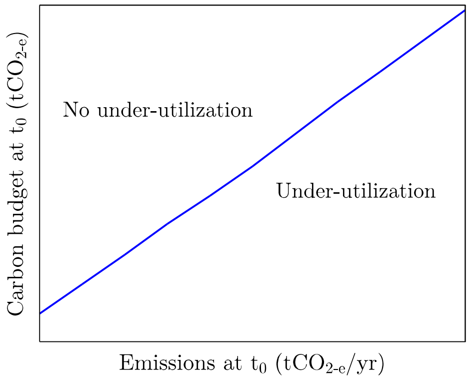

Figure 2: Depending on initial emissions (i.e. initial brown capital $k_{b,0}$) and on the concentration ceiling ($\bar{m}$), brown capital is underutilized or not in the first-best optimum.

the underutilization of brown capital happens if initial brown capital is high (right end of the x-axis) and/or if the ceiling is stringent (lower part of the y-axis).

During the first phase when the irreversibility cost is positive, if the irreversibility cost is too high, part of polluting capital can become obsolete and both the acquisition price (ν_t) and rental rate ($R_{p,t}$) of polluting capital fall down to zero.[12] While all polluting capital looses value in the short-run, this obsolete capital becomes stranded as it is early-scrapped to increase short-term abatement.

In this section, we have found that under irreversible investment, society has to live with past mistakes for a while, once it realizes it has been on a non-optimal growth path. A way to limit the associated irreversibility cost is to give up part of installed polluting capital in order to reduce emissions faster (thereby creating stranded assets). In the next section, we find that investment-based policies reduce emissions without affecting existing polluting capital, and therefore increase the social cost of GHG abatement.

4. Investment-based instruments

Current climate mitigation policies rarely include carbon prices and rely instead on investment-based instruments such as energy efficiency standards or fiscal incentives for green investment (as feebates, which impose additional fees on polluting capital and rebates for clean capital). These instruments redirect investment towards clean capital but have no effect on the use of existing capital.

In this section, we investigate the optimal transition to a clean-capital economy using investment-based instruments. We find that (i) they are less efficient than the first-best carbon tax in terms of inter-temporal welfare maximization, (ii) they allow reaching the same steady state, and (iii) they induce a full utiliza-

[12]The strictly positive marginal productivity of utilized polluting capital is transferred to households through the tax revenue $\tau_t\ G$.

tion of polluting capital in the short run, thereby reducing short-term income losses.

A way to trigger the transition to a clean economy is to differentiate investment costs with feebate programs, i.e. fiscal incentives that include subsidies on clean investment ($\theta_{c,t} > 0$) and taxes on polluting investment ($\theta_{p,t} > 0$). With such a feebate program, π_t, the flow of firms' net receipt at time t is equal to:

$$\pi_t = F(q_{p,t}, q_{c,t}) - (\lambda_t - \theta_{c,t})\, i_{c,t} - (\lambda_t + \theta_{p,t})\, i_{p,t} \tag{32}$$

Where λ_t is the price of investments. The optimal values of $\theta_{c,t}$ and $\theta_{p,t}$ can be obtained with a maximization of social welfare given the ceiling constraint. The same steady state as in the social optimum is reached (at a date $t_{ss,2}$ which is different than $t_{ss,1}$ in general). On the steady-state the optimal value of $\theta_{c,t} + \theta_{p,t}$ is equal to the first-best carbon tax multiplied by the marginal emissions of polluting capital:[13]

$$\forall t \geq t_{ss,2},\ \theta_{c,t} + \theta_{p,t} = \tau_{t,1} \cdot G \tag{33}$$

Investment-based instruments however induce a different short-term transition. Over the short-run, investment in polluting capital stops. However, as firms do not pay carbon emissions directly, it is never optimal to underutilize polluting capital (Appendix D). As a consequence, short-term output may be higher than in the first-best strategy:[14]

Proposition 6. *With a feebate program, short-term output is equal or higher than in the first-best solution with a carbon price.*

Proof. The first-best carbon price may induce underutilization of polluting capital in the short-run ($q_{p,1,t} < k_{p,t}$). In the second-best solution capital is not underused ($q_{p,2,t} = k_{p,t}$). As both strategies start with a phase during which investment is nill, installed capital is identical with both policies in an interval $(t_0, \tilde{t})$. During this interval, utilized polluting capital, hence output, is higher in the second-best strategy. $\square$

Similarly to the carbon price, investment-based instruments differentiate the marginal productivities of capital (Appendix D):

$$\partial_{q_p} F = \partial_{k_c} F - \underbrace{\frac{1}{\lambda_t}\Big((\rho+\delta)\psi_t - \dot{\psi_t}\Big)}_{p_t} + \underbrace{\frac{1}{\lambda_t}\Big((\delta+\rho)(\theta_{c,t}+\theta_{p,t}) - (\dot{\theta_{c,t}} + \dot{\theta_{p,t}})\Big)}_{\tau_{t,2}\ G} \tag{34}$$

where p_t is the irreversibility cost. In this second-best setting the shadow price of carbon $\tau_{t,2}$ is still equal to a technical abatement cost plus the irreversibility

[13] Note that the same outcome can be reached using taxes on polluting investment alone or subsidies to clean investment alone, since what matters is the sum of the tax plus the subsidy. A tax and a subsidy lead to different transfers in the society, which can play a key role on the acceptability of a particular environmental policy (e.g Sterner and Höglund Isaksson, 2006; Fischer, 2008)

[14] Analytically the effect on consumption is ambiguous because it involves the offsetting impacts from a substitution effect and an income effect: short-term output is higher, but investments in clean capital may also increase since the saving rate is endogenous.

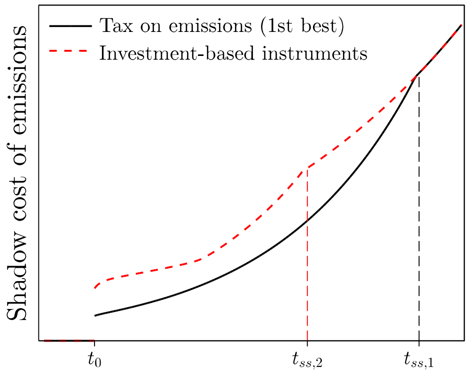

Figure 3: The shadow price of emissions (or carbon price) is higher with investment-based instruments than with a carbon price.

cost:

$$\underbrace{\tau_{t,2}}_{\text{economic cost}} = \underbrace{\frac{\partial_{q_p} F - \partial_{k_c} F}{G}}_{\text{technical cost}} + \underbrace{\frac{p}{G}}_{\text{irreversibility cost}} \tag{35}$$

The irreversibility cost p is no longer bounded by the marginal productivity of clean capital but by the shadow price of carbon $\tau_{t,2}$ (Appendix D). Indeed, preventing underutilization is like refusing to recognize that past accumulation of polluting capital was a mistake. When society keeps using obsolete polluting capital instead of early-scrapping it, the irreversibility cost can be as high as the cost of the carbon emissions that installed brown capital produces.

The short-term utilization of obsolete polluting capital leads to a different shadow price of carbon than in the first-best case with a carbon tax:

$$\forall t \in I_{u,1},\ \tau_{t,2} - \tau_{t,1} = \frac{1}{G}\left[\partial_{q_p} F(q_{p,2} = k_p) - \partial_{q_p} F(q_{p,1} < k_p) + \alpha_t\right] \tag{36}$$

Where $I_{u,1}$ is the interval during which capital is underutilized on the first-best pathway and $\alpha_t = p_t - \partial_{k_c} F$ is the extra cost associated to the utilization of obsolete polluting capital on the second-best pathway. This extra-cost can be interpreted as a temporary subsidy on the utilization of polluting capital in the welfare-maximization framework (Appendix E).

Figure 3 compares the shadow prices of carbon with the first and second-best policies. Investment-based policies generate a higher emissions shadow price than the carbon tax alone, however the dynamics of capital accumulation mean that the social cost of abatement cannot be translated into consumption losses in a trivial way (see also Vogt-Schilb et al., 2014). Even if investment-based instruments set a higher emissions shadow cost at each time t (Fig. 3), output is higher over the short-run (Prop. 6, Fig. 4).

Investment-based policies only differ temporarily from the first-best pathway, in a way that smooths the transition costs: they decrease efforts in the short-run (Prop. 6), leave them unchanged in the long-run (eq. 33), and (therefore) increase efforts in the medium-run (Fig. 4). Moreover, feebate programs induce a different intra-generational distribution of abatement efforts from the carbon

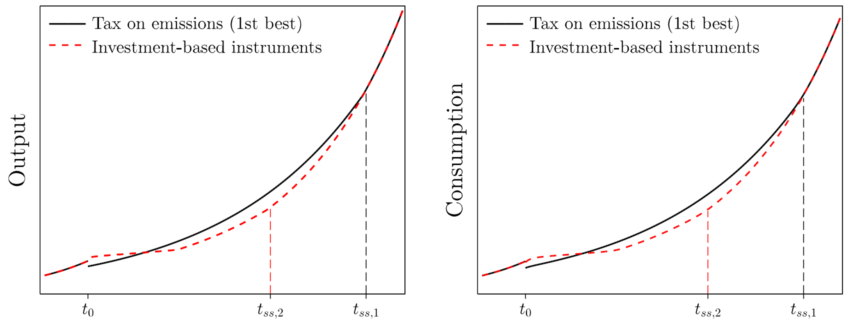

Figure 4: On the left, output y in the two cases. In the short-run output is lower in the first-best case because of the adjustment of polluting capital utilization. On the right, consumption c is higher in the second-best case because of a higher output y. t_{ss} is the date at which the steady state is reached, it is reached sooner in the second-best case ($t_{ss,2} < t_{ss,1}$).

tax, since they avoid stranded assets. By preventing new investment in polluting capital, they even increase the value of existing polluting assets:

Proposition 7. *With a feebate program, the market price of polluting capital is initially higher than the price of clean capital, and than the marginal utility of consumption.*

PROOF. First-order conditions for firms' receipt maximization give:

$$\nu_t = \chi_t + \theta_{c,t} + \theta_{p,t} - \psi_t \tag{37}$$

where ν_t is the price of polluting capital and χ_t is the price of clean capital. The policy creates a scarcity effect on polluting capital, that increases its price while the irreversibility constraint reduces its price in the short-run. Appendix D shows that $\theta_{c,t} + \theta_{p,t} - \psi_t \geq 0$. □

Investment-based instruments are not limited to feebates and may include performance standards for new capital.[15] Such performance standards include for instance existing energy-efficiency standards for new vehicles, buildings, and appliances.

Proposition 8. *In our model the optimal feebate program is equivalent to the optimal performance standard on new capital: (1) they induce the same investment and output pathways and (2) they have the same impact on the price of polluting capital.*

PROOF. Appendix F.

As with feebates, performance standards induce a full utilization of existing polluting capital in the short-run and redirect investments towards clean capital.

[15] Investment costs can also be differentiated using financial markets, as proposed by Rozenberg et al. (2013).

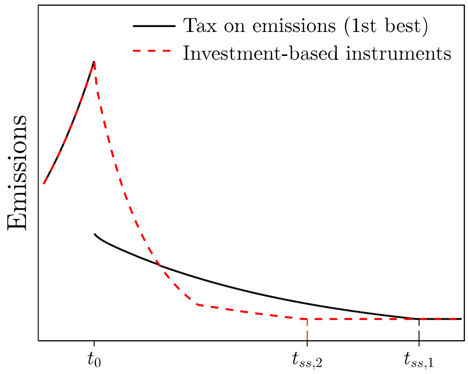

Figure 5: GHG emissions in the two cases. The carbon price induces spare polluting capital and thus reduces carbon emissions faster in the short-run.

They also create scarcity on existing polluting capital and therefore increase the price of polluting capital with regards to clean capital on secondary markets.[16]

5. Timing of action and carbon-intensive lock-in

The utilization of investment-based instruments is limited by their slowness in reducing emissions. Indeed, since they maintain a full utilization of polluting capital in the short-run, investment-based policies result in higher short-term emissions than the carbon tax (Prop. 6 and Fig. 5) and might not be sufficient for stringent climate objectives as regards to past capital accumulation.

Figure 6 proposes a visualization of this issue. Starting from low polluting capital stocks (thus low emissions), a carbon tax does not lead to underutilization of polluting capital and reaching the climate target is possible and optimal without a downward step in income. In this case, the carbon price consistent with the climate target leads to the exact same pathway as investment-based policies. This is a situation of "flexibility" in which a country can chose a polluting or a clean development path at low cost, using either a carbon price or investment-based instruments.

But as long as climate policies are absent (or very lax), the economy accumulates polluting capital, making GHG emissions grow and reducing the residual carbon budget for a given climate target (the arrow "conventional growth").

At one point, the threshold when the marginal productivity of polluting capital is lower than the optimal carbon price is crossed (see eq. 31), meaning that polluting capital should be underutilized and output reduced. From there, a carbon price becomes more difficult to implement because of political-economy constraints. But the alternative option of using investment-based instruments is available, leading to higher inter-temporal costs but no immediate drop in income. There is a window of opportunity, during which alternative investment-based instruments may induce a smooth and acceptable transition to a low-carbon economy.

[16]In this model the capital lifetime is endogenous and therefore people cannot extend the lifetime of their polluting capital.

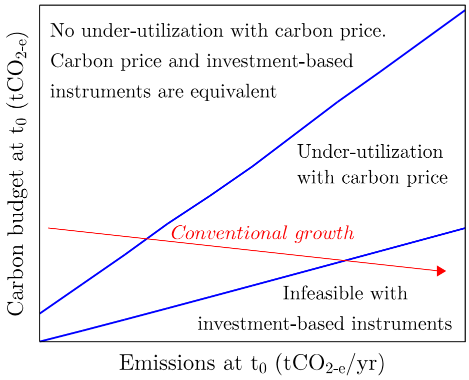

Figure 6: Depending on initial emissions (i.e. initial polluting capital k_{b,t_0}) and on the carbon budget $(\bar{m} - m_{t_0})$, the carbon tax and investment-based instruments can lead to different or similar outcomes (for a given set of parameters, and in particular ρ and δ). If the carbon budget is too stringent, such that waiting for polluting capital depreciation is not sufficient, the investment-based instruments cannot be used. If the carbon budget is not stringent, there is no underutilization of polluting capital in the first-best optimum with the carbon tax and investment-based instruments are equivalent. While the economy is on the laissez-faire growth path (red arrow), polluting capital accumulates and the carbon budget is reduced for a given climate objective.

If this occasion is missed (right hand side, Fig. 6), it becomes impossible to reach the climate target without underutilization of polluting capital and investment-based options are not available any more (if the climate objective is not revised). In this last area, not only the economic cost of reaching the climate target is higher, but the political economy also creates a carbon lock-in: the only option to reach the climate target requires early-scrapping and thus has a significant short-term cost, making it more difficult to implement successfully a climate policy consistent with the target.

The zone in which polluting capital must be underutilized to remain below the ceiling depends on the capital depreciation rate δ, the GHG dissipation rate ε, initial GHG concentration m_0 and initial polluting capital k_0. The lower blue line in Fig. 6 is expressed analytically in Appendix G and can be approximated by:

$$\bar{m} < m_0 + \frac{G\, k_0}{\delta}$$

According to Davis et al. (2010), the level of existing polluting infrastructure in 2010 is still low enough to achieve the 2°C target without underutilizing polluting capital, suggesting that the global economy is not in this last region yet. They find that if existing energy infrastructure was used for its normal life span and no new polluting devices were built, future warming would be less than 0.7°C. Yet, reaching the 2°C target might imply to stop investing in polluting capital very soon, which depends on our ability to overcome infrastructural inertia and develop clean energy and transport services (Davis et al., 2010; Guivarch and Hallegatte, 2011). Note that Davis et al. (2010) do not discuss whether the least-cost policy would lead to underutilization, that is whether we are in the top or the middle triangle in Fig. 6.

6. Discussion

Choosing the best instrument in terms of welfare results in choosing the lowest social cost of abatement but not the highest consumption at each time t. There is a trade-off between efficiency (maximum intertemporal welfare), intergenerational equity (distribution of efforts over time) and implementation obstacles (political economy). The carbon tax is the best tool to maximize discounted welfare, but public policy is especially difficult in contexts where costs are immediate, concentrated and visible, while benefits are invisible (avoided damages) and diffuse over time and over citizens (Olson, 1977). Policy-makers may use other criteria than social welfare maximization to choose the policy to implement (Beltratti et al., 1994; Chichilnisky et al., 1995).

One possible reason why investment-based instruments are preferred by policy-makers is that they give the owners of existing polluting capital some time to adapt to the new economic conditions – without carrying a loss for past decisions – and prevent capital underutilization. Indeed, underutilization of capital may appear as a waste of resources, results in an output drop and creates unemployment. Also, the owners of obsolete polluting capital and the workers whose jobs depend on this capital can be strong opponents to climate policies. Governments may thus be captured by the owners of polluting capital (Laffont and Tirole, 1991) who claim compensations because they invested under pre-existing rules and will own stranded assets. Finally, governments may also be captured by individuals who have different time preferences from the social planner's. Indeed, time preference heterogeneity makes it unappealing for some people to pay now for remote future benefits. This is even more so because future generations are likely to be richer and the ones benefiting from reduced climate change damages. Since investment-based strategies postpone mitigation efforts to the medium-run, they would be preferred by people with high discount rates.

Investment-based instruments therefore ease the political economy of the transition to clean capital. While the outcome of such instruments is lower in terms of discounted intertemporal welfare, they have the potential to tackle both the effectiveness (they trigger a transition to clean capital) and the equity (they compensate losers) functions of a climate mitigation policy, as well as inter-generational distributional issues.

Our analysis is incomplete and further analyses of the distributional impacts of mitigation instruments could model capital retrofit (an intermediary solution between investing in new clean capital and early-scrapping existing polluting capital) or learning-by-doing and knowledge spillovers (which would improve the productivity of clean capital). We also omitted to consider cases with myopic agents or limited ability to commit. Nevertheless, our results suggest that investment-based instruments respond to a political acceptability issue as regards to climate mitigation policies.

Acknowledgements

The authors thank Patrice Dumas, Marianne Fay, Louis-Gaetan Giraudet, Baptiste Perissin-Fabert and the participants of the 2014 AEA conference for their useful comments on a previous version of this article. All remaining errors are the authors'. We acknowledge financial support from the Institut pour la

Mobilité Durable. The views expressed in this paper are the sole responsibility of the authors. They do not necessarily reflect the views of the World Bank, its executive directors, or the countries they represent.

References

Acemoglu, D., Aghion, P., Bursztyn, L., Hemous, D., 2012. The environment and directed technical change. American Economic Review 102 (1), 131–166.

Allcott, H., 2013. The welfare effects of misperceived product costs: Data and calibrations from the automobile market. American Economic Journal: Economic Policy 5 (3), 30–66.

Allen, M. R., Frame, D. J., Huntingford, C., Jones, C. D., Lowe, J. A., Meinshausen, M., Meinshausen, N., 2009. Warming caused by cumulative carbon emissions towards the trillionth tonne. Nature 458 (7242), 1163–1166.

Ambrosi, P., Hourcade, J., Hallegatte, S., Lecocq, F., Dumas, P., Ha Duong, M., 2003. Optimal control models and elicitation of attitudes towards climate damages. Environmental Modeling and Assessment 8 (3), 133–147.

Anderson, S. T., Parry, I. W. H., Sallee, J. M., Fischer, C., 2011. Automobile fuel economy standards: Impacts, efficiency, and alternatives. Review of Environmental Economics and Policy 5 (1), 89–108.

Arltesou, A., 1999. Models of energy use: Potty-potty versus putty-clay. The American Economic Review 89 (4), 1028–1043.

Arrow, K. J., Cline, W. R., Maler, K. G., Munasinghe, M., Squitieri, R., Stiglitz, J. E., 1996. Intertemporal equity, discounting, and economic efficiency. Cambridge University Press.

Arrow, K. J., Kurz, M., 1970. Optimal growth with irreversible investment in a Ramsey model. Econometrica 38 (2), 331–344.

Beltratti, A., Chichilnisky, G., Heal, G., 1994. The environment and the long run: a comparison of different criteria. Ricerche Economiche 48 (4), 319–340.

Carbon Tracker, 2013. Unburnable carbon 2013: Wasted capital and stranded assets. Tech. rep., The Grantham Research Institute on Climate Change and the Environment of LSE.

Chichilnisky, G., Heal, G., Beltratti, A., 1995. The green golden rule. Economics Letters 49 (2), 175–179.

Davis, S. J., Caldeira, K., Matthews, H. D., 2010. Future CO2 emissions and climate change from existing energy infrastructure. Science 329 (5997), 1330 –1333.

Fischer, C., 2008. Comparing flexibility mechanisms for fuel economy standards. Energy Policy 36 (8), 3116–3124.

Fischer, C., Withagen, C., Toman, M., 2004. Optimal investment in clean production capacity. Environmental and Resource Economics 28 (3), 325–345.

Gerlagh, R., Kverndokk, S., Rosendahl, K. E., 2009. Optimal timing of climate change policy: Interaction between carbon taxes and innovation externalities. Environmental and Resource Economics 43 (3), 369–390.

Goulder, L. H., Hafstead, M. A. C., Dworsky, M., 2010. Impacts of alternative emissions allowance allocation methods under a federal cap-and-trade program. Journal of Environmental Economics and Management 60 (3), 161–181.

Goulder, L. H., Mathai, K., 2000. Optimal CO2 abatement in the presence of induced technological change. Journal of Environmental Economics and Management 39 (1), 1–38.

Goulder, L. H., Parry, I. W. H., 2008. Instrument choice in environmental policy. Review of Environmental Economics and Policy 2 (2), 152–174.

Grimaud, A., Lafforgue, G., Magné, B., 2011. Climate change mitigation options and directed technical change: A decentralized equilibrium analysis. Resource and Energy Economics 33 (4), 938–962.

Guivarch, C., Hallegatte, S., 2011. Existing infrastructure and the 2C target. Climatic Change 109 (3-4), 801–805.

Helm, D., Hepburn, C., Mash, R., 2003. Credible carbon policy. Oxford Review of Economic Policy 19 (3), 438–450.

IEA, 2014. Policies and measures database. International energy agency. URL http://iea.org/policiesandmeasures

Jenkins, J. D., In press. Political economy constraints on carbon pricing policies: What are the implications for economic efficiency, environmental efficacy, and climate policy design? Energy Policy.

Jorgenson, D., 1967. The theory of investment behavior. In: Determinants of investment behavior. NBER.

Kanbur, R., 2010. Macro crisis and targeting transfers to the poor. In: Globalization and Growth: Implications for a Post-Crisis World. Michael Spence, Danny Leipziger, Commission on Growth and Development, The World Bank, Washington DC, p. 342.

Kydland, F. E., Prescott, E. C., 1977. Rules rather than discretion: The inconsistency of optimal plans. The Journal of Political Economy 85 (3), 473–491.

Laffont, J.-J., Tirole, J., 1991. The politics of government decision-making: A theory of regulatory capture. The Quarterly Journal of Economics 106 (4), 1089–1127.

Matthews, H. D., Gillett, N. P., Stott, P. A., Zickfeld, K., 2009. The proportionality of global warming to cumulative carbon emissions. Nature 459 (7248), 829–832.

Nordhaus, W. D., 1991. To slow or not to slow: The economics of the greenhouse effect. The Economic Journal 101 (407), 920–937.

Olson, M., 1977. The logic of collective action: public goods and the theory of groups. Harvard University Press.

Parry, I. W. H., Evans, D., Oates, W. E., In press. Are energy efficiency standards justified? Journal of Environmental Economics and Management.

Ploeg, F. V. D., Withagen, C., 1991. Pollution control and the ramsey problem. Environmental and Resource Economics 1 (2), 215–236.

Rozenberg, J., Hallegatte, S., Perrissin-Fabert, B., Hourcade, J.-C., 2013. Funding low-carbon investments in the absence of a carbon tax. Climate Policy 13 (1), 134–141.

Slechten, A., 2013. Intertemporal links in cap-and-trade schemes. Journal of Environmental Economics and Management 66 (2), 319–336.

Sterner, T., Höglund Isaksson, L., 2006. Refunded emission payments theory, distribution of costs, and swedish experience of NOx abatement. Ecological Economics 57 (1), 93–106.

Tsvetanov, T., Segerson, K., 2013. Re-evaluating the role of energy efficiency standards: A behavioral economics approach. Journal of Environmental Economics and Management 66 (2), 347–363.

Vogt-Schilb, A., Meunier, G., Hallegatte, S., 2014. Optimal timing, cost and sectoral distribution of emission reductions: abatement cost curves vs. abatement investment. ASSA, Philadelphia.

Wei, C., 2003. Energy, the stock market, and the putty-clay investment model. The American Economic Review 93 (1), 311–323.

Williams, R. C., 2011. Setting the initial time-profile of climate policy: The economics of environmental policy phase-ins. NBER, 245–254.

Appendix A. Maximization of the household's utility

The household maximizes their inter temporal utility (eq. 2) given the motion law of wealth (eq. 1). The present value Hamiltonian is:

$$H_h(c_t, a_t) = e^{-\rho t} \cdot \{u(c_t) + \lambda_t [r_t \cdot a_t + y_t - c_t]\} \tag{A.1}$$

where λ_t is the shadow cost of investment in assets at time t. The first order conditions for a maximum of W are:

$$\forall t,\ \partial_c H_h = 0 \Rightarrow \lambda_t = u'(c_t) \tag{A.2}$$

$$\forall t,\ \partial_a H_h + \frac{\partial(e^{-\rho t}\lambda_t)}{\partial t} = 0 \Rightarrow \dot{\lambda}_t = (\rho - r_t)\lambda_t \tag{A.3}$$

The doted variables represent temporal derivatives. Differentiating eq. A.2 with respect to time and substitute for λ from eq. A.3, yields the Euler equation:

$$\frac{\dot{c}_t}{c_t} = \frac{-u'(c_t)}{c_t \cdot u''(c_t)} \cdot (r_t - \rho) \tag{A.4}$$

Appendix B. Social optimum (section 3)

The present value Hamiltonian associated to the maximization of social welfare (16) is:

$$
\begin{aligned}
H_t = e^{-\rho t} \cdot \{ & u(c_t) + \lambda_t[F(q_p, k_c) - c_t - i_{p,t} - i_{c,t}] + \nu_t[i_{p,t} - \delta k_{p,t}] \\
& + \chi_t[i_{c,t} - \delta k_{c,t}] - \mu_t \cdot [G\, q_{p,t} - \varepsilon m_t] + \phi_t \cdot [\bar{m} - m_t] \\
& + \psi_t \cdot i_{p,t} + \beta_t[k_{p,t} - q_{p,t}]\}
\end{aligned} \tag{B.1}
$$

All multipliers are positive.
The complementary slackness conditions are:

$$\forall t,\ \psi_t \geq 0 \text{ and } \psi_t \cdot i_{p,t} = 0 \tag{B.2}$$

$$\forall t,\ \beta_t \geq 0 \text{ and } \beta_t \cdot (k_{p,t} - q_{p,t}) = 0 \tag{B.3}$$

$$\forall t,\ \phi_t \geq 0 \text{ and } \phi_t \cdot (\bar{m} - m_t) = 0 \tag{B.4}$$

Appendix B.1. First order conditions

First order conditions give:

$$\frac{\partial H_t}{\partial c_t} = 0 \Rightarrow \qquad u'(c_t) = \lambda_t \tag{B.5}$$

$$\frac{\partial H_t}{\partial i_{p,t}} = 0 \Rightarrow \qquad \lambda_t = \nu_t + \psi_t$$

$$\frac{\partial H_t}{\partial i_{c,t}} = 0 \Rightarrow \qquad \lambda_t = \chi_t$$

$$\frac{\partial H_t}{\partial k_{p,t}} = -\frac{\partial(e^{-\rho t}\nu_t)}{\partial t} \Rightarrow \qquad -\nu_t\delta + \beta_t = -\dot{\nu}_t + \rho\nu_t$$

$$\frac{\partial H_t}{\partial k_{c,t}} = -\frac{\partial(e^{-\rho t}\chi_t)}{\partial t} \Rightarrow \quad \lambda_t \partial_{k_c} F(k_{p,t}, k_{c,t}) - \chi_t\delta = -\dot{\chi}_t + \rho\chi_t$$

$$\frac{\partial H_t}{\partial q_{p,t}} = 0 \Rightarrow \qquad \lambda_t \partial_{q_p} F(q_{p,t}, k_{c,t}) - \mu_t \cdot G = \beta_t$$

$$\frac{\partial H_t}{\partial m_t} = \frac{\partial(e^{-\rho t}\mu_t)}{\partial t} \Rightarrow \qquad -\phi_t + \varepsilon\mu_t = \dot{\mu}_t - \rho\mu_t \tag{B.6}$$

Appendix B.2. Equilibrium on the capital market and interest rate: proof of proposition 3

If we differentiate eq. B.5 with respect to time and substitute λ_t and $\dot{\lambda}_t$, we can write:

$$\frac{c_t \cdot u''(c_t)}{u'(c_t)} \cdot \frac{\dot{c}_t}{c_t} = (\rho + \delta - R_{c,t}) \tag{B.7}$$

As in the laissez-faire equilibrium (eq. A.4), the interest rate r_t that ensures households are indifferent between consumption and investment is thus given by:

$$r_t := R_{c,t} - \delta \tag{B.8}$$

Appendix B.3. Carbon price

Eq. B.6 gives the evolution of μ_t. Using $\dot{\mu_t} = (\dot{\lambda_t \tau_t} + \lambda_t \dot{\tau_t})$ (from eq. 21), eq. B.5, eq. B.7 and eq. B.8 yields:

$$\dot{\tau_t} = \tau_t[\varepsilon + r_t] - \frac{\phi_t}{\lambda_t}$$

We call t_{ss} the date at which GHG concentration reaches the ceiling:

$$\forall t \geq t_{ss},\ m_t = \bar{m}$$

During the steady state, $\dot{m_t} = 0 \implies G\,q_{p,t} = \varepsilon\,\bar{m}$ (eq. 11). On the long run, installed capital is not underused, polluting installed capital is thus constant at $k_{p,t} = \bar{m}\,\varepsilon/G$ during the steady state.

Before t_{ss}, $\phi_t = 0$ (B.4). The carbon price thus exponentially grows at the endogenous interest rate plus the dissipation rate of GHG until the ceiling is reached:

$$\dot{\tau_t} = \tau_t[\varepsilon + r_t] \tag{B.9}$$

These dynamics may be interpreted as a generalized Hotelling rule applied to clean air: along the optimal pathway, and before the ceiling is reached, the discounted abatement costs are constant over time. The appropriate discount rate is $r_t + \varepsilon$, to take into account the natural decay of GHG in the atmosphere.

Appendix B.4. The irreversibility constraint is binding in the short run : proof of proposition 4

A binding GHG ceiling is imposed at t_0. Before that, the economy was in the competitive equilibrium, such that clean and polluting capital have the same marginal productivity and installed capital is fully used (Proposition 1):

$$\lim_{t \to t_0^-} q_{p,t} = k_{p,t} \tag{B.10}$$

$$\lim_{t \to t_0^-} \partial_{q_p} F(q_{p,t}, q_{c,t}) = \partial_{k_c} F(q_{p,t}, q_{c,t}) \tag{B.11}$$

We use a proof by contradiction to show that at t_0^+ (when the constraint is internalized) the irreversibility condition is necessarily binding. Suppose that the transition starts with a phase when the irreversibility constraint is not binding, i.e. $\psi_t = 0$. This would lead to (Propositions 2 and 3):

$$\lim_{t \to t_0^+} \partial_{q_p} F(q_{p,t}, q_{c,t}) = \partial_{k_c} F(q_{p,t}, q_{c,t}) + \tau_{t_0} \cdot G \tag{B.12}$$

Besides, investment means that capital is a continuous function of time:

$$\lim_{t \to t_0^+} q_{p,t} = k_{p,t} \tag{B.13}$$

If the GHG ceiling is binding then $\tau_{t_0} > 0$ (eq. B.9). So from eq. B.11 and eq. B.12:

$$\lim_{t \to t_0^+} \partial_{q_p} F(q_{p,t}, q_{c,t}) \neq \lim_{t \to t_0^+} \partial_{q_p} F(q_{p,t}, q_{c,t}) \tag{B.14}$$

$\partial_{q_p} F$ is a continuous function of $q_{p,t}$ so eq. B.14 implies that $\lim_{t \to t_0^+} q_{p,t} \neq \lim_{t \to t_0^+} q_{p,t}$, which is incompatible with eq. B.10 and eq. B.13.

□

Appendix C. Decentralized equilibrium with a tax on emissions

In a decentralized economy, it is possible to trigger the same outcome as in the social optimum with a lump-sum tax applied to carbon emissions. In this case, the firm's flow of profit at time t is given by:

$$\Pi_t = F(q_{p,t}, k_{c,t}) - R_{c,t} \cdot k_{c,t} - R_{p,t} \cdot k_{p,t} - \tau_t \, G \, q_{p,t} \tag{C.1}$$

With $R_{p,t}$ and $R_{c,t}$ the rental prices of polluting and clean capacities respectively, and τ_t the carbon tax. The tax is redistributed through the assets equation:

$$\dot{a}_t = r_t \cdot a_t + y_t - c_t + \tau_t \, G \, q_{p,t} \tag{C.2}$$

The Lagrangian corresponding to the firm's maximization program is:

$$L(t) = \Pi_t + \beta_t(k_{p,t} - q_{p,t}) + \gamma_t(k_{c,t} - q_{c,t}) \tag{C.3}$$

First order conditions are:

$$\begin{aligned}
\partial_{q_g} L = 0 \Rightarrow \qquad & \partial_{q_c} F(q_{p,t}, q_{c,t}) = \gamma_t && \text{(C.4)}\\
\partial_{q_b} L = 0 \Rightarrow \qquad & \partial_{q_p} F(q_{p,t}, q_{c,t}) = \beta_t + \tau_t \cdot G && \text{(C.5)}\\
\partial_{k_g} L = 0 \Rightarrow \qquad & \gamma_t = R_{c,t} && \text{(C.6)}\\
\partial_{k_b} L = 0 \Rightarrow \qquad & \beta_t = R_{p,t} && \text{(C.7)}
\end{aligned}$$

For all t,

$$\gamma_t \geq 0 \text{ and } \gamma_t \cdot (k_{c,t} - q_{c,t}) = 0$$
$$\beta_t \geq 0 \text{ and } \beta_t \cdot (k_{p,t} - q_{p,t}) = 0$$

(complementary slackness conditions).
With eq. C.4 we have $\gamma_t = \partial_{q_c} F(q_{p,t}, q_{c,t}) > 0$, so $q_{c,t} = k_{c,t}$ for all t.
The combination of eq. C.4 and eq. C.6 gives

$$\partial_{k_c} F(q_{p,t}, k_{c,t}) = R_{c,t}$$

Combining eq. C.5 and eq. C.7, we find

$$\partial_{q_p} F(q_{p,t}, k_{c,t}) = R_{p,t} + \tau_t \cdot G \tag{C.8}$$

In the equilibrium, the rental price of clean capacities is equal to the interest rate (plus delta): $R_{c,t} = r_t + \delta$, because clean capacities and loans are perfect substitutes as assets for households. When the irreversibility constraint is not binding (see eq. 6), and in particular on the balanced growth path, the rental rate of polluting capacities is equal to the interest rate as well and $R_{p,t} = R_{c,t} = r_t + \delta$.

However, when the carbon price in implemented at t_0, the irreversibility constraint is binding (4). In this case, since the use of polluting capacities suddenly becomes too expensive, the rental rate of polluting capacities is endogenously reduced. As a consequence of a lower rate of return for owners of polluting capital, households stop investing in polluting capacities. If the carbon tax is very high, the rental rate of polluting capacities can even become nil and polluting capacities may be under-utilized.

Appendix D. Firms' maximization problem with differentiation of investment costs

The present value Hamiltonian associated to the firm's maximization program is:

$$
\begin{aligned}
H_t = e^{-\rho t} \cdot \{ & F(q_{p,t}, q_{c,t}) - (\lambda_t - \theta_{c,t})\, i_{c,t} - (\lambda_t + \theta_{p,t})\, i_{p,t} \\
& + \nu_t [i_{p,t} - \delta k_{p,t}] + \chi_t [i_{c,t} - \delta k_{c,t}] \\
& + \psi_t \cdot i_{p,t} + \beta_t [k_{p,t} - q_{p,t}] \}
\end{aligned}
$$

First order conditions give:

$$
\begin{aligned}
\frac{\partial H_t}{\partial i_{p,t}} = 0 \Rightarrow & \qquad \lambda_t + \theta_{p,t} = \nu_t + \psi_t \\
\frac{\partial H_t}{\partial i_{c,t}} = 0 \Rightarrow & \qquad \lambda_t - \theta_{c,t} = \chi_t \\
\frac{\partial H_t}{\partial k_{p,t}} = -\frac{\partial (e^{-\rho t}\nu_t)}{\partial t} \Rightarrow & \qquad -\nu_t \delta + \beta_t = -\dot{\nu}_t + \rho \nu_t \\
\frac{\partial H_t}{\partial k_{c,t}} = -\frac{\partial (e^{-\rho t}\chi_t)}{\partial t} \Rightarrow & \qquad \rho \chi_t - \dot{\chi}_t = \lambda_t \partial_{k_c} F(k_{p,t}, k_{c,t}) - \chi_t \delta \\
\frac{\partial H_t}{\partial q_{p,t}} = 0 \Rightarrow & \qquad \lambda_t \partial_{q_p} F(q_{p,t}, k_{c,t}) = \beta_t
\end{aligned}
\tag{D.1}
$$

The complementary slackness condition $\forall t,\ \beta_t [k_{p,t} - q_{p,t}] = 0$ combined with equation D.1 gives that — if F satisfies the Inada conditions — capital is never underused with investment-based instruments $\forall t, k_{p,t} = q_{p,t}$.

FOCs can be reduced to:

$$\nu_t + \psi_t = \chi_t + \theta_{c,t} + \theta_{p,t} \tag{D.2}$$

$$\partial_{k_c} F = \frac{1}{\lambda}((\delta + \rho)\chi_t - \dot{\chi}_t) \tag{D.3}$$

$$\partial_{q_p} F = \frac{1}{\lambda}((\delta + \rho)\nu_t - \dot{\nu}_t) \tag{D.4}$$

We thus obtain

$$\partial_{q_p} F = \partial_{k_c} F + \underbrace{\frac{1}{\lambda_t}\Big((\delta + \rho)(\theta_{c,t} + \theta_{p,t}) - (\dot{\theta_{c,t}} + \dot{\theta_{p,t}})\Big)}_{\theta_t} - \underbrace{\frac{1}{\lambda_t}\Big((\rho + \delta)\psi_t - \dot{\psi}_t\Big)}_{p_t} \tag{D.5}$$

With p_t the irreversibility cost and θ_t a positive term that depends on $(\theta_{c,t} + \theta_{p,t})$.

Equation D.5 is similar to eq. E.4 with $\theta_t = \tau_{t,2}\ G$, where $\tau_{t,2}$ is the shadow price of carbon. In the optimal pathway with a full-utilization of capital, θ_t is therefore equal to the shadow price of carbon (multiplied by G).

In this setting under-utilizing polluting capital is never optimal because firms do not pay carbon emissions directly. Instead, investment in polluting capital is more expensive that investment in clean capital and over the short-run, as in the social optimum the economy does not invest in new polluting capital. Once polluting capital has depreciated to a level compatible with the GHG ceiling, polluting investments become profitable and start again.

The policy creates a scarcity effect on polluting capital, that increases its price ($\theta_{c,t} + \theta_{p,t}$, eq. D.2) while the irreversibility constraint reduces its price in the short-run (ψ_t, eq. D.2).

Along the optimal transition to the new long-term steady state,

$$\begin{aligned} &\partial_{k_c} F \leq \partial_{q_p} F \\ \Leftrightarrow & p_t \leq \theta_t (= \tau_{t,2}) \\ \Leftrightarrow & \psi_t \leq \theta_{c,t} + \theta_{p,t} \end{aligned} \tag{D.6}$$

so that the price of pre-existing polluting capital is higher than that of clean capital is the short-run.

In the steady state, the irreversibility cost is nill ($p = 0$) and the marginal productivity of polluting capital is equal to that of clean capital plus θ_t. The same steady state as in the social optimum is reached and the optimal value of θ_t is equal to the first-best carbon tax multiplied by the marginal emissions of polluting capital:

$$\forall t \geq t_{ss}, \ \theta_t = \tau_t \cdot G$$

with t_{ss} the date at which the steady state is reached.

With investment-based instruments, the shadow price of emissions $\tau_{t,2}$ is still equal to a technical abatement cost plus the irreversibility cost:

$$\underbrace{\tau_{t,2}}_{\text{economic cost}} = \underbrace{\frac{\partial_{q_p} F - \partial_{k_c} F}{G}}_{\text{technical cost}} + \underbrace{\frac{p}{G}}_{\text{irreversibility cost}} \tag{D.7}$$

$$\text{with } p \in [0, \ \tau_{t,2}]$$

The irreversibility cost p is now bounded by the shadow carbon price $\tau_{t,2}$ (eq. D.6). One interpretation is that preventing under-utilization is like refusing to recognize that past accumulation of polluting capital may have been a mistake. By doing so, the irreversibility cost can be has high as the cost of the GHG emissions that installed brown capital produces.

Appendix E. Maximization of social welfare with full utilization constraint: temporary subsidy on existing polluting capital

The same outcome as with feebates or standards can be reached with the same social planner program as in Appendix B and a full-utilization constraint:

$$\max_{c,i,k} \int_0^\infty e^{-\rho t} \cdot u(c_t) \, dt \tag{E.1}$$

$$\begin{aligned} \text{subject to } & F(q_p, k_c) - c_t - i_{p,t} - i_{c,t} = 0 && (\lambda_t) \\ & \dot{k}_{p,t} = i_{p,t} - \delta k_{p,t} && (\nu_t) \\ & \dot{k}_{c,t} = i_{c,t} - \delta k_{c,t} && (\chi_t) \\ & \dot{m}_t = G \, q_{p,t} - \varepsilon m_t && (\mu_t) \\ & m_t \leq \bar{m} && (\phi_t) \\ & i_{p,t} \geq 0 && (\psi_t) \\ & q_{p,t} \leq k_{p,t} && (\beta_t) \\ & q_{p,t} = k_{p,t} && (\alpha_t) \end{aligned}$$

The present value Hamiltonian associated to the maximization of social welfare is:

$$\begin{aligned}H_t = e^{-\rho t} \cdot \{u(c_t) + \lambda_t[F(q_p, k_c) - c_t - i_{p,t} - i_{c,t}] + \nu_t[i_{p,t} - \delta k_{p,t}] \\ +\chi_t[i_{c,t} - \delta k_{c,t}] - \mu_t \cdot [G\, q_{p,t} - \varepsilon m_t] + \phi_t \cdot [\bar{m} - m_t] \\ + \psi_t \cdot i_{p,t} + \beta_t[k_{p,t} - q_{p,t}] + \alpha_t[q_{p,t} - k_{p,t}]\}\end{aligned}$$

All multipliers are positive.
Equations 19 and 20 become:

$$\begin{aligned}\beta_t - \alpha_t &= \frac{1}{\lambda}((\delta + \rho)\nu_t - \dot{\nu}_t) \\ \partial_{q_p} F &= \beta_t - \alpha_t + \tau_t \cdot G\end{aligned}$$

The rental price of polluting capital is therefore equal to $\beta_t - \alpha_t$. The condition on the marginal productivity of polluting capital becomes:

$$\partial_{q_p} F = \beta_t - \alpha_t + \tau_t \cdot G \tag{E.2}$$

Note that due to complementary slackness conditions, if $\beta_t > 0$ then $\alpha_t = 0$ and if $\alpha_t > 0$ then $\beta_t = 0$. In the first phase when polluting investment is nil, if the carbon tax is higher than the marginal productivity of the last unit of polluting capital, the value of polluting capital is nil, $\beta_t = 0$ and the equation becomes:

$$\partial_{q_p} F = -\alpha_t + \tau_t \cdot G \tag{E.3}$$

α_t is a subsidy to the utilization of polluting capital.
Similarly to the first-best pathway, the marginal productivities are differentiated as follows:

$$\begin{aligned}&\partial_{q_p} F = \partial_{k_c} F - p_t + \tau_t\, G \qquad \text{(E.4)} \\ &0 < p_t < \tau_t\, G\end{aligned}$$

With the irreversibility cost $p_t > 0$ during the first phase and $p_t = 0$ when polluting capital reaches a sustainable level.

In the long run when $i_b > 0$ the equilibrium is equivalent to the social optimum. In the short run when $i_b = 0$, $\psi_t > 0$ and $R_{p,t} < R_{c,t}$, except that in this case $R_{p,t}$ becomes negative if the carbon price is higher than the marginal productivity of the last unit of polluting capital (expressed in output per emissions). Thus polluting capital is always fully-utilized.

This instrument leads to the same investments and output as the differentiation of investment costs or standards, however it is not perfectly equivalent. Indeed, the carbon tax also affects polluting capital on the secondary markets, thus the price of polluting capital decreases in the short-run. Conversely, with taxes on investments or standards on investments, polluting capital becomes scarce and so its price increase on the secondary market.

An instrument perfectly equivalent to the tax plus subsidy would be to differentiate capital costs, that is to tax both polluting investment and exchanges on the secondary market.

Appendix F. Investment regulation (performance standards)

Another equivalent possibility is to regulate polluting investment through efficiency standards. In particular, the most polluting investments can be forbidden. Here, we crudely impose polluting investments to be nil until polluting capital has depreciated to a level allowing to reach the carbon ceiling without under-utilizing polluting capital.

We come back to the social planner's program (beginning of section 3) and remove the concentration and ceiling constraints (eq. 11 and eq. 15). We can also remove the irreversibility constraint (eq. 6) which will not be binding in this case. Instead, we add a polluting investment constraint that forces $i_{p,t}$ to be equal to a standard at each point in time, and we call σ_t its Lagrangian multiplier:

$$\forall t,\ i_{p,t} = sd_t \qquad (\sigma_t) \tag{F.1}$$

The standard sd_t can be optimally set to equal polluting investments found in the previous section and the next section. Basically, $sd_t = 0$ until polluting capacities have depreciated to a level compatible with the ceiling. The present value Hamiltonian associated to the maximization of social welfare is:

$$\begin{aligned} H_t = e^{-\rho t} \cdot \{ & u(c_t) + \lambda_t[F(q_p, k_c) - c_t - i_{p,t} - i_{c,t}] + \nu_t[i_{p,t} - \delta k_{p,t}] \\ & + \chi_t[i_{c,t} - \delta k_{c,t}] + \sigma_t \cdot (sd_t - i_{p,t}) + \beta_t[k_{p,t} - q_{p,t}]\} \end{aligned} \tag{F.2}$$

λ_t is the current value shadow price of income. ν_t and χ_t are the current shadow values of investments in polluting and clean capital.
First order conditions can be reduced to the following equations:

$$u'(c_t) = \lambda_t = \nu_t - \sigma_t = \chi_t \tag{F.3}$$

$$\lambda_t \partial_{k_c} F = (\delta + \rho)\chi_t - \dot{\chi}_t \tag{F.4}$$

$$\lambda_t \partial_{q_p} F = \beta_t \tag{F.5}$$

$$\beta_t = (\delta + \rho)\nu_t - \dot{\nu}_t \tag{F.6}$$

Here, σ_t is equivalent to $(\theta_{c,t} + \theta_{p,t} - \psi_t)$ in the previous section.

The maximization of intertemporal welfare results in the same equations as in the previous sections:

$$\begin{aligned} & R_{p,t} = R_{c,t} + n_t \\ & \text{with } n_t = \frac{1}{\lambda_t}\left((\rho + \delta)\sigma_t - \dot{\sigma_t}\right) \end{aligned} \tag{F.7}$$

This equation is equivalent to Eq. D.5, with $n_t = \theta_t - p_t$. The variable n_t is positive, which means that the rental price of polluting capacities is higher than the interest rate. Indeed, as with the differentiation of investment costs the polluting investment standard creates a scarcity effect on polluting capital, which becomes more expensive than clean capital.

This instrument must be thought of as temporary, since once polluting capital has depreciated to a sustainable level, a carbon price can be implemented without inducing under-utilization of polluting capital, and thus becomes politically acceptable. Investment regulation can be compared with existing efficiency standards on cars or electric plants, that forbid the construction of the most polluting kinds of polluting capital.

Appendix G. Second-best infeasibility zone

This zone defines the cases when the ceiling is reached before polluting capacities have depreciated to a sustainable level. If no investment is made in polluting capacities, we have:

$$k_{p,t} = k_0 \, e^{-\delta t}$$

Therefore, the stock of pollution follows this dynamic:

$$\dot{m} = k_0 \, e^{-\delta t} - \varepsilon \, m$$

The solution to this differential equation is:

$$m_t = -\frac{G \, k_0}{\delta - \varepsilon} \, e^{-\delta t} + \left(m_0 + \frac{G \, k_0}{\delta - \varepsilon} \right) \, e^{-\varepsilon t}$$

This function first increases to a maximum $m_{max} = \frac{G \, k_0}{\delta} e^{-\delta t}$ and then decreases. The maximum date is

$$t_{max} = -\frac{1}{\delta} \, \ln(\frac{m_{max} \, \varepsilon}{G \, k_0})$$

The expression of m at the maximum date gives the limit of the infeasibility zone if $m_{max} = \bar{m}$:

$$\bar{m} = -\frac{G \, k_0}{\delta - \varepsilon} \, e^{\ln(\frac{\bar{m} \, \varepsilon}{G \, k_0})} + \left(m_0 + \frac{G \, k_0}{\delta - \varepsilon} \right) \, e^{\frac{\varepsilon}{\delta} \ln(\frac{\bar{m} \, \varepsilon}{G \, k_0})}$$

This can be rewritten:

$$\bar{m} = \left[\left(m_0 + \frac{G \, k_0}{\delta - \varepsilon} \right) \left(\frac{\varepsilon}{G \, k_0} \right)^{\frac{\varepsilon}{\delta}} \left(\frac{\delta - \varepsilon}{\delta} \right) \right]^{\frac{\delta}{\delta - \epsilon}}$$

The "clean incentives infeasibility zone" depends on the capital depreciation rate, the GHG dissipation rate, initial GHG concentration and initial polluting capacities.