Instrument Choice and Stranded Assets in the Transition to Clean Capital

Julie Rozenberg[1] , Adrien Vogt-Schilb[2], and Stephane Hallegatte[3]

Abstract

To mitigate climate change, some governments opt for instruments focused on new investment, like mandates, feebates, or performance standards instead of carbon prices that would affect existing capital as well. We compare these policies in a Ramsey model with clean and polluting capital, irreversible investment and a climate constraint. Alternative instruments imply different transitions to the same balanced growth path. The optimal carbon price minimizes the discounted social cost of the transition to clean capital, but may create political economy issues by prompting premature retirement of existing polluting capacities and creating concentrated private costs in the form of stranded assets. Second-best mandates or feebates on new capital lead to higher social costs but smooth abatement costs over individuals and time, do not result in premature retirement, and avoid stranded assets. A phased-in carbon price can avoid premature retirement but still result in stranded assets, that is in a drop of wealth for the owners of polluting capital. We discuss a potential trade-off between political feasibility and cost-effectiveness of environmental policies.

JEL: L50, O33, O44, Q52, Q54, Q58.
Keywords: political economy; public acceptability; policy instrument; climate change; intergenerational equity; energy efficiency standards; mothballing; early-scrapping; creative destruction.

This version: May 2018

1. Introduction

For the past centuries, economic growth has involved the accumulation of fossil-fueled capital, such as coal power plants and gasoline-fueled cars, which release greenhouse gases (GHG) to the atmosphere. Later, science has established that this release of GHG, if sustained over centuries, would result in dangerous climate change. To stabilize the climate and manage disruption and damages, economies now have to reduce emissions to near-zero levels (IPCC, 2014; Fay et al., 2015). Doing so implies a transition from production based on polluting capital to production based on clean, carbon-neutral capital. In principle, the optimal policy to enforce such a transition is to use a carbon price (Pigou, 1932; Nordhaus, 1991; Pearce, 1991), imposed through a carbon market or, perhaps preferably (Goulder and Schein, 2013), a carbon tax. Combined with targeted innovation policies, a carbon price could redirect investment away from polluting and towards clean capital at a relatively low cost (IPCC, 2014).

However, such carbon prices may create immediate *stranded assets*, that is assets that suffer from premature write downs, devaluations, or that become a financial liability. The words *stranded assets* are used in the literature on climate change to describe various things (Caldecott et al., 2016): assets that are lost because of the impact of climate change itself, fossil fuel resources that cannot be burnt into the atmosphere if a given climate target is to be reached (Asheim, 2012), also called *unburnable carbon* (McGlade and Ekins, 2015), and physical man-made capital which book value has to be written off, or that has to be retired early because it was not built taking into account the new climate policies, such as coal power plants that become unprofitable after a carbon price is implemented (Guivarch and Hood, 2010).

This paper focuses on stranded man-made capital. The lifetime of equipment in the transportation, building, industry, and energy sectors range from about a decade — in the case of a car — to half a century — for power plants — or even centuries — for city shapes and transportation systems (Davis et al., 2010; Guivarch and Hallegatte, 2011; Sachs et al., 2014). Older vintages of this capital have been built before climate change was identified as a serious threat, and many of existing vintages have been built to be used for several decades into the future (Davis and Socolow, 2014; Rozenberg et al., 2015; Shearer et al., 2017; Pfeiffer et al., 2018). A sudden change in prices induced by environmental taxes may result in a drop of the net income those asset generate, causing them to lose a significant share of their book value prematurely. For instance, a taxi company would write the book value of its fleet off if gasoline prices increase significantly. At worse, carbon prices could make carbon-intensive assets become entirely unprofitable. For instance, Johnson et al. (2015) estimate that a carbon price consistent with the 2°C target will result in the premature retirement of at least 165 billion US dollars worth of coal power plants worldwide.

In a recent speech, the Governor of the Bank of England has expressed concern that the magnitude of those stranded assets could be a threat for the stability of the financial system (Carney, 2016). In addition, it has been argued that stranded assets may create political economy issues and obstruct the implementation of substantial carbon prices (Jenkins, 2014; Bertram et al., 2015; Vogt-Schilb and Hallegatte, 2017). Indeed, stranded assets are a visible loss of wealth concentrated in a few vested interests, such as the coal industry, who may oppose the reform — and in some cases may even have the power to veto it (Olson, 1977; Trebilcock, 2014). Furthermore, a literature on public attitudes towards environmental taxes suggests that there is an aversion to carbon taxes partly driven by (i) the perception that they are inefficient, absent clean alternatives to polluting activities; and (ii) the perception that they are unfair, as their cost is perceived to fall disproportionately on a few actors (e.g., Dresner et al., 2006; Kallbekken and Sælen, 2011; Harrison and Peet, 2012). Stranded assets are related to both issues, as they are a symptom of limited availability of clean alternatives in the short term, and they translate into immediate costs concentrated on a few actors.

This paper uses a simple model to investigate how alternative policy instruments may reduce stranded assets and premature retirement of polluting capital during the transition to clean capital. We focus on the effect of instruments such as mandates for new power plants, buildings and appliances, moratoriums or bans of new coal power plants, *feebate* programs that tax new energy-inefficient equipment and subsidize new energy-efficient equipment, energy efficiency standards on new equipment, or subsidized loans and tax breaks for energy efficiency investment. All these instruments are similar in that they redirect private investment away from polluting capital and toward clean capital without directly affecting the owners of the existing stock of polluting capital, for instance without providing incentive to drive less or operate existing gas power plants instead of existing coal power plants.

We analyse how using carbon prices, mandates, feebates, or performance standards leads to different costs and dynamics of the transition from polluting to clean capital, with a particular focus on premature retirement and revenues from existing capital. We compare the first-best carbon price, designed to reduce the discounted cost of the transition, and second-best feebates or mandates designed to minimize costs while avoiding premature retirement. A motivation for comparing these types of instruments is that while some governments have enacted carbon prices (World Bank, 2016), most existing emission-reduction policies regulate only new investment (IEA, 2016; IPCC, 2014, p. 28). In addition, phased-in carbon prices have been proposed as a way to reduce adjustment costs (Williams, 2012); and actual implementation of carbon prices, for instance in British Columbia and France, have been phased-in progressively. We thus also look at the second-best phased-in carbon price schedule designed to minimize costs while avoiding premature retirement.

We use a Ramsey model with two types of capital (as in Acemoglu et al., 2012): *polluting* capital, which creates GHG emissions, and *clean* capital, which does not. We disregard knowledge spillovers and we model the climate change constraint as a GHG concentration ceiling. Investment is assumed irreversible, as in (Arrow and Kurz, 1970): existing polluting capital cannot be converted back into consumption or transformed into clean capital. We however allow for under-utilization of existing polluting capital, a feature that is generally omitted in multi-sector growth models. Under-utilization means that emission-reduction effort can be divided between two qualitatively different channels: (i) long-term abatement through accumulation of clean capital instead of polluting capital (e.g. agents buy electric cars instead of gasoline-fueled cars); and (ii) immediate abatement through the underutilization or early decommissioning of polluting capital (e.g. agents drive less or scrap their gasoline cars).

We find that all policy instruments lead to the same long-term growth path, in which most installed capital is clean and carbon concentration is maintained at its maximum acceptable value. The optimal carbon price and the second-best phased-in carbon price, feebates, and mandates however induce different short-term pathways in terms of emissions and costs, and in particular different levels and distribution of costs.

Unsurprisingly, the carbon price minimizes the total discounted cost of the climate change policy. Under a carbon price, investment is redirected towards clean capital until polluting capital has depreciated to a level compatible with the concentration ceiling. In addition, the carbon price creates stranded assets. In particular, part of the existing polluting capital is under-utilized or retired prematurely if climate policies are stringent enough — that is, if the carbon price is larger than the marginal productivity of polluting capital over its carbon intensity. Such outcomes are part of the least-cost strategy, because under-utilization or early-scrapping of polluting capital reduces carbon emissions from excessive legacy polluting assets. But this strategy sets a disproportionate cost on the owners of polluting capital:[1] premature retirement reduces their revenues to zero while polluting capacities are under-used. Even if it does not cause premature retirement, the carbon price creates stranded assets: it ties a new cost to the utilization of existing polluting capital and thus decreases its net value.

In contrast, mandates, feebates, or standards on new investment do not tie a new cost to the utilization of existing capital. These instruments thus avoid stranded assets and their extreme form, the premature retirement of polluting capital. The second-best phased-in carbon price designed to avoid early retirement results in the same transition and same social costs, but it fails to avoid stranded assets: it may adjust at the level that cancels all revenues for the owners of current capital — such that they are indifferent between renting out their assets or scraping them.

Mandates, feebates, and phased-in carbon prices are less efficient that the first-best carbon price. They create higher social costs: society keeps using obsolete polluting capacities until the end of their lifetime instead of scrapping them — as if refusing to recognize that past accumulation of polluting capital was a mistake. We track analytically the social cost resulting from past excessive investment in dirty capacity (directly linked to the assumption that investment is irreversible), and we call it *legacy cost*. With phased-in carbon prices, as with optimal carbon prices, those legacy costs fall directly on the owners of stranded

---

[1] By disproportionate we mean in relation to their proportion in the population.

assets in the form of reduced rental rates. But feebates and mandates do not create stranded assets. These instruments can even increase the relative value of existing polluting capacity (and thus create windfall profits for their owners), as they create a scarcity on the supply of polluting capacities without tying a cost to their utilization.

These results suggest that policy makers face a choice between (1) a higher intertemporal welfare with the optimal carbon price; (2) less immediate costs with a phased-in carbon prices; or (3) less immediate costs and no stranded assets with feebates, mandates, or standards.

Finally, another important difference between the optimal carbon price and the other instruments is their mere *efficacy*. As they do not lead to decommissioning any polluting capital, second-best feebates, standards and phased-in carbon prices reduce emissions slower than the optimal carbon price, and cannot achieve too stringent GHG concentration targets — while the optimal carbon price could reduce emissions arbitrarily quickly. Empirical evidence suggests that it could still be technically possible to reach the 2°C target while avoinding stranded assets. For instance, Davis et al. (2010) estimate that emissions embedded in existing long-lived capital and infrastructure in 2010 committed us to a warming of about 1.3°C. However, findings by Rogelj et al. (2013), Johnson et al. (2015), and Iyer et al. (2015) suggest that the least-cost pathway toward a 2°C-compliant economy does involve stranding assets. Governments willing to limit global warming below 2°C might still have a choice between first-best carbon prices and second-best feebates or standards, and this choice would imply a trade-off between minimizing discounted costs and avoiding stranded assets.

The remainder of the paper is structured as follows. Section 2 details our contribution to specific branches of the literature. Section 3 presents the basic model and considers the *laissez-faire* equilibrium. In section 4 we consider a social planner who maximizes utility under a climate constraint, and show how its strategy can be decentralized with a carbon price. In section 5, we consider a modified social planner program where premature retirement is to be avoided, and we look at how second-best mandates, feebates, or a phased-in carbon price can decentralize that constrained strategy. In section 6, we study the timing issues and risks of lock-in when premature retirement is avoided. Section 7 concludes.

2. Contribution to the Literature

This paper relates to several branches of the literature.

First, the literature on instrument choice for environmental policy has established that the carbon price is the most efficient instrument (Pigou, 1932; Goulder and Parry, 2008; Fischer and Newell, 2008). For instance, the extensive literature on CAFE standards concludes that they are less efficient than a carbon tax, because they do not provide incentive to reduce emissions from the existing fleet (Austin and Dinan, 2005), may even create a rebound effect, worsening the effect of unaddressed externalities such as congestion or emission of local pollutants (Anderson et al., 2011), slow down capital turnover, reducing the speed at which the new, energy-efficient cars enter the fleet (Jacobsen and van Benthem, 2015), and can hurt poorer households by increasing the price of new and used cars (Jacobsen, 2013). All these important considerations are left out of our model.

In terms of distributional impacts, many ex-ante studies focus on their incidence on household of different income categories (e.g., Rausch et al., 2010; Fullerton et al., 2012; Coady et al., 2015; Borenstein and Davis, 2015) or different generations (e.g., Karp and Rezai, 2012). Other papers explore how different policies set costs on different sectors of the economy. One is Fullerton and Heutel (2010), who find in a two-sector static model that the additional welfare cost of performance standards, compared to that generated by a carbon price, is not supported by the dirty sector, but spread over the clean one—Giraudet and Quirion (2008) had reached a similar conclusion in the case of policies that promote energy efficiency. We expand this literature by comparing alternative instruments in the dynamic context of the transition to clean capital, showing that the optimal carbon prices and a second-best phased-in carbon prices, feebates, or standards lead not only to different distribution of costs between sectors but also over time.

By modelling investment and production decisions separately, we also show that feebates and standards are not entirely equivalent to a carbon tax plus a production subsidy — as previous research focused on efficiency impacts has found (e.g., Fischer and Newell, 2008; Holland et al., 2009). Both options may lead to the same social cost and result in the same investment and production decisions, but their incidence are different. Feebates and standards operate by influencing investment decisions and do not directly reduce income for the owners of existing polluting capital, while a carbon-tax-plus-subsidy scheme operates by influencing production and does reduce the value of the existing stock of polluting capital.

Further, Goulder et al. (2010) have studied how carbon markets can be designed to compensate firms for stranded assets, and find that under a cap-and-trade system, the owners of polluting firms may be fully compensated if a fraction of emissions allowances are grandfathered for free—making all permits free is likely to result in substantial windfalls profits for the owners of polluting capital (e.g., Sijm et al., 2006). Goulder and Schein (2013) note that the same result can be obtained with carefully-designed exemptions under a carbon tax.[2] We expand this literature by comparing carbon prices with alternative instruments: instead of offering compensation to the owners, standards and feebates avoid stranded assets and premature retirement and all their potential impacts

---

[2] It is well established that a potential advantage of carbon pricing schemes over regulations — not captured in our model — is that the remaining revenues from carbon pricing can be used to mitigate policy costs by reducing other distortive fiscal policies (Bovenberg and Goulder, 1996; Parry and Bento, 2000; Metcalf, 2014; Rausch and Reilly, 2015; Siegmeier et al., 2018). Standards do not have this feature, but feebates may result in net fiscal benefits (or net fiscal costs, or be revenue-neutral), and phased-in carbon prices do raise revenue.

in the first place. The second-best phased-in carbon price can avoid premature retirement, but not stranded assets, as it adjusts at the maximum level that reduces revenues from polluting capital to zero, which makes little difference from the point of view of their owners. (In principle, the government could still use the revenues anticipated from the carbon phase-in, or grandfather the rights in a phased-in market, to compensate the owners of polluting capital for stranded assets.)

Finally, our paper relates to the literature that studies the transition to a clean economy through the lens of the directed technical change theory (e.g. Gerlagh et al., 2009; Kalkuhl et al., 2012; André and Smulders, 2014). This literature studies the policy mix to tackle both the climate change externality and sector-specific knowledge accumulation and spillovers. One finding highlighted by Kverndokk and Rosendahl (2007), Grimaud and Lafforgue (2008) and Acemoglu et al. (2012) is that, in the short term, the least-cost policy relies relatively more on research subsidies in the clean sector than on carbon prices. The reason is that the most powerful lever to reduce GHG emissions is to encourage investment into a structural transformation of the economy over the long term, not to distort production decisions in the short term. Here we disregard knowledge accumulation and focus instead on another feature of Schumpeterian creative destruction: the retirement of capital prematurely made obsolete by the accumulation of new generations of physical capital. Our findings suggest that feebates and mandates can trigger structural change while avoiding immediate disruption of the old sector.

3. Decentralized model and laissez-faire equilibrium

3.1. Model

Capitalist household. A representative household aims at maximizing its intertemporal utility:

$$\int_0^\infty e^{-\rho t} \cdot u(c_t)\, dt \tag{1}$$

where $\rho > 0$ is its pure rate of time preference, u is a classic isoelastic utility function, and c_t is consumption at time t.

The households owns physical capital of two sorts: clean capital $k_{c,t}$ and polluting capital $k_{p,t}$. It makes money by renting out a portion $q_{c,t}$ and $q_{p,t}$ of available clean and polluting capacities $k_{c,t}$ and $k_{p,t}$ to producers at the respective rates $R_{p,t}$ and $R_{c,t}$ that they take as given. They use this revenue to purchases goods for consumption c_t or invest in capacities:

$$R_{c,t} \cdot q_{c,t} + R_{p,t} \cdot q_{p,t} = c_t + i_{p,t} + i_{c,t} \tag{2}$$

$$q_{p,t} \leq k_{p,t} \tag{3}$$

$$q_{c,t} \leq k_{c,t} \tag{4}$$

In the remainder of this paper, q will be called *utilized capital* and k *installed capital* or *capacity*. The underutilization of polluting capacities can be optimal when facing a constraint on GHG emissions. For instance, all coal plants in the economy can be operated part-time, or some of them can be shut down, if the utilization of the whole capital stock is conflicting with the climate objective. Both cases are captured in aggregate with $q_{p,t} < k_{p,t}$. In this paper, it turns out that underutilization of clean capital is never optimal, so we omit the difference between $q_{c,t}$ and $k_{c,t}$ for short in the remainder of the paper.

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

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

$$\dot{k}_{c,t} = i_{c,t} - \delta\, k_{c,t} \tag{6}$$

The doted variables represent temporal derivatives.

Investment is assumed to be irreversible:[3]

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

$$i_{c,t} \geq 0 \tag{8}$$

This means that for instance, a coal plant cannot be turned into a wind turbine, and only disappears through natural depreciation.

Arrow and Kurz (1970) study the consequence of irreversible investment in a Ramsey model with one type of (clean) capital. They show that if the initial stock of clean capital is higher than its long-term value, the solution to the Ramsey problem with irreversible investment is to let clean capital depreciate until it has reached its long-term optimal level. Otherwise, the solution is the same as in the classic Ramsey model (where the irreversibility of investment is omitted): investment is set at the level where the value of capital and the value of consumption coincide. In this paper, we assume that clean capital is not initially over-abundant, and disregard for short the implications of the irreversibility of clean investment. We thus ignore (8) for the rest of the paper.

Producer. At time t, a representative producer produces one final good y_t, using polluting capital $q_{p,t}$ and clean capital $k_{c,t}$.

$$y_t = F(A_t, q_{p,t}, k_{c,t}) \tag{9}$$

where A_t is the exogenous total factor productivity, assumed to increase exponentially over time, and F is a constant-returns production function, assumed to satisfy the Inada conditions and be smooth enough.[4]

The producer rents the capital $q_{p,t}$ and $k_{c,t}$ from the household at the respective rental rates $R_{p,t}$ and $R_{c,t}$, taken as given. The producer maximizes the following profit:

$$\max_{q} \Pi(q_{p,t}, k_{c,t}) = F(A_t, q_{p,t}, k_{c,t}) - R_{c,t} \cdot k_{c,t} - R_{p,t} \cdot q_{p,t} \tag{10}$$

---

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

[4] A_t can be understood as effective technologically-augmented labor (Barro and Sala-i-Martin, 2003), since we omit labor in this paper.

The environment. Polluting capital used at time t emits greenhouse gases e_t:

$$e_t = G \times q_{p,t}$$

Where G is a strictly positive constant representing the carbon intensity of polluting capital. Our modelling of emissions and production is not perfect. For instance, we leave capital retrofit, which would not only increase the stock of clean capital but also reduce the stock or the carbon intensity of dirty capital, for further research.

Nonetheless, our model captures parsimoniously the functioning of the most energy-intensive sectors of the economy, which are responsible for the bulk of carbon emissions: power generation, transportation, and buildings (light and air conditioning). In these sectors, greenhouse gas emissions depend to a significant extent on the technology *embedded* in existing capital (for instance a given type of coal power plant, light bulb, or car), and how much the capital is used (how many hours per year a light bulb is on, how much electricity is generated from the plant, how many kilometres are travelled by the car). We also omit labour in the production function, essentially assuming that substituting labour for capital (e.g. drivers for taxis, operators for coal plants, or domestic workers for light bulbs) is not a prominent option to reduce GHG emissions from existing capital.

GHG atmospheric concentration m_t increases with emissions, and decreases with a dissipation rate ε:

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

The dissipation rate makes it possible to maintain a small stock of polluting capital once the transition is over, which simplifies exposition, but the policy conclusions hold if we assume $\varepsilon = 0$.

3.2. Laissez-Faire Equilibrium

Household. To set up a useful baseline for the rest of the paper, this section considers the laissez-faire equilibrium, when the household disregards the environment. We also disregard momentarily the effects of the irreversibility constraint on polluting capital (7) — without a carbon ceiling, the two types of capital do not have any practical difference and the analysis by Arrow and Kurz (1970) can be generalized easily: the irreversibility constraint is binding only if one type of capital is above its long-term path (which does not happen in a simple growth framework), in that case the optimal solution would be to let capital depreciate until it reaches back its equilibrium value. We keep a more detailed study of irreversibility constraint for the more interesting case of overabundant polluting capital when an environmental policy is implemented in the next section.

Here, the household's problem is to maximize discounted welfare under budget and capacity constraints:

$$\max_{c,i,q} \int_0^\infty e^{-\rho t} \cdot u(c_t)\, dt \tag{12}$$

$$\begin{aligned} \text{subject to } & R_{c,t} \cdot k_{c,t} + R_{p,t} \cdot q_{p,t} = c_t + i_{p,t} + i_{c,t} && (\lambda_t) \\ & \dot{k}_{p,t} = i_{p,t} - \delta k_{p,t} && (\nu_t) \\ & \dot{k}_{c,t} = i_{c,t} - \delta k_{c,t} && (\chi_t) \\ & q_{p,t} \leq k_{p,t} && (\beta_t) \end{aligned}$$

where ρ is the rate of time preference. We indicated in parentheses the co-state variables and Lagrangian multipliers (chosen such that they are positive): among them, λ_t is the shadow value of income (used as numeraire), ν_t and χ_t are the shadow value of new polluting and clean capital, and the Lagrange multiplier β_t interprets as the shadow cost of the polluting capacity constraint.

The FOCs can be written as (Appendix A):

$$u'(c_t) = \lambda_t = \nu_t = \chi_t \tag{13}$$

$$R_{c,t} = \frac{1}{\lambda_t}[(\delta + \rho)\chi_t - \dot{\chi}_t] \tag{14}$$

$$R_{p,t} = \frac{1}{\lambda_t}[(\delta + \rho)\nu_t - \dot{\nu}_t] = \frac{\beta_t}{\lambda_t} \tag{15}$$

Equation (15) combined with the complementary slackness condition,

$$\beta_t \geq 0 \text{ and } \beta_t \cdot (k_{p,t} - q_{p,t}) = 0 \tag{16}$$

implies that all polluting capital is rented out if the rental price for polluting capital is positive $R_{p,t} > 0$. In the laissez-faire equilibrium, this is always the case because, with Inada-compliant production function, the producer is always willing to pay a positive rent on polluting capital (see eq. 19 below).

Equation (13) implies that in the laissez-faire equilibrium, the household chooses consumption and investment such that the value of polluting and clean capital are equal, and are both equal to the value of consumption (the numeraire).

As explained by Jorgenson (1967), the relationship between the rental costs ($R_{c,t}$, $R_{p,t}$) and the prices of new capital (χ_t, ν_t) captured by equations (14) and (15) ensures agents would be indifferent between buying and renting capital, given the depreciation rate δ, the pure preference for present ρ, and the future price of capital (implied by $\dot{\chi}_t$ and $\dot{\nu}_t$). Alternatively, integrating these differential equations shows that the shadow values of capacities equal the net present value of future rents received by a deprecating capacity, plus a salvage value:

$$\forall T > t, \quad \chi_t = \int_t^T e^{-(\rho+\delta)(\tau-t)} \lambda_\tau R_{c,\tau} d\tau + e^{-(\rho+\delta)T} \chi_T \tag{17}$$

$$\nu_t = \int_t^T e^{-(\rho+\delta)(\tau-t)} \lambda_\tau R_{p,\tau} d\tau + e^{-(\rho+\delta)T} \nu_T \tag{18}$$

Producers. To maximize profits, producers simply need to observe the rental rates $R_{p,t}, R_{c,t}$ and rent capital up to the point where marginal returns equal the respective rental rates:

$$\frac{\partial \Pi}{\partial q_p} = 0 \implies \quad \frac{\partial F}{\partial q_p}(q_{p,t}, k_{c,t}) = R_{p,t} \tag{19}$$

$$\frac{\partial \Pi}{\partial k_c} = 0 \implies \quad \frac{\partial F}{\partial k_c}(q_{p,t}, k_{c,t}) = R_{c,t} \tag{20}$$

The markets for physical capacities clear if the household wants to rent out the same quantity of capital as the producer wants to use. In that case, the rental rates and the quantities that appear in the FOCs of the household and those that appear in the FOCs of the firm coincide. Combining the FOCs (13), (14), (15), (19) and (20) one finds that:

Lemma 1. *In the laissez-faire equilibrium, if the irreversibility constraint is not binding, the rental rates of polluting and clean capital are equal, and the marginal productivity of clean and polluting capital are also equal:*

$$R_{c,t} = R_{p,t} \tag{21}$$

$$\frac{\partial F}{\partial q_p}(q_{p,t}, k_{c,t}) = \frac{\partial F}{\partial k_c}(q_{p,t}, k_{c,t}) \tag{22}$$

This familiar result translates the well-known equi-marginal principle.

This result is simple but provides a useful benchmark.[5] In the following sections, we compare the effects of irreversible investment, different social constraints and policy instruments to this benchmark.

4. Minimizing Inter-temporal Costs under a Ceiling Constraint

In this section, we assume that the economy is on the laissez-faire equilibrium as described in Lemma 1, and at a time t_0 the social planner suddenly decides to internalize an environmental target. This can be interpreted as an economy that unexpectedly discovers an environmental externality and immediately introduces a policy to internalize it, or a situation in which agents have failed to anticipate the introduction or the stringency of an environmental policy.

We consider the socially-optimal transition to a clean economy and how it can be decentralized with a carbon tax.

4.1. The Social Planner's Optimum

In this section, we assume a social planner is willing to maintain the atmospheric concentration of carbon m_t below a given ceiling $\bar{m}$ (à la Chakravorty et al., 2006):

$$m_t \leq \bar{m} \tag{23}$$

---

[5] By ignoring the lower bound on investment, Lemma 1 is focusing on the singular solution of the problem. The next sections discuss bang-singular solutions explicitly.

Ceiling constraints are consistent with a cautious cost-effectiveness approach (Manne and Richels, 1992; Ambrosi et al., 2003; Weitzman, 2012). The ceiling 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 2°C target in the Paris Agreement, or any other temperature target designed to hedge society against catastrophic climate change (IPCC, 2014).

The social planner directly choses consumption, investment, and production to maximize inter-temporal utility given the constraints set by the economy budget, the capital motion law, investment irreversibility, capacity constraints, and the GHG ceiling. The social planner program is:

$$\max_{c,i,q} \int_0^\infty e^{-\rho t} \cdot u(c_t)\, dt \tag{24}$$

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

where μ_t is the shadow price of carbon, expressed in terms of utility at time t, ψ_t is the Lagrange multiplier associated with the irreversibility of investment, the Lagrange multiplier β_t interprets as the shadow cost of the polluting capacity constraint, and ϕ_t is the Lagrange multiplier associated with the carbon ceiling.

The Lagrangian associated to the constrained maximization of social welfare can be found in Appendix B.1. The first-order conditions of the social planner's problem boil down to:

$$u'(c_t) = \lambda_t = \nu_t + \psi_t = \chi_t \tag{25}$$

$$\frac{\partial F}{\partial k_c} = \frac{1}{\lambda_t}\left((\delta + \rho)\chi_t - \dot{\chi}_t\right) \tag{26}$$

$$\beta_t = \left((\delta + \rho)\nu_t - \dot{\nu}_t\right) \tag{27}$$

$$\frac{\partial F}{\partial q_p} = \frac{\beta_t}{\lambda_t} + \tau_t \cdot G \tag{28}$$

Where τ_t is the social cost of carbon expressed in dollars per ton:

$$\tau_t := \frac{\mu_t}{\lambda_t} \tag{29}$$

The complementary slackness conditions are:

$$\forall t,\ \psi_t \geq 0 \text{ and } \psi_t \cdot i_{p,t} = 0 \tag{30}$$

$$\forall t,\ \beta_t \geq 0 \text{ and } \beta_t \cdot (k_{p,t} - q_{p,t}) = 0 \tag{31}$$

$$\forall t,\ \phi_t \geq 0 \text{ and } \phi_t \cdot (\bar{m} - m_t) = 0 \tag{32}$$

In the right hand side of equations (26) and (27) we recognize the values of the rental rates of clean and polluting capital found in the previous section, that we thus call the implicit rental cost of capital and denote $R^i_{c,t}$ and $R^i_{p,t}$:

$$R^i_{c,t} := \frac{1}{\lambda_t}[(\delta + \rho)\chi_t - \dot{\chi}_t] \tag{33}$$

$$R^i_{p,t} := \frac{1}{\lambda_t}[(\delta + \rho)\nu_t - \dot{\nu}_t] \tag{34}$$

The system tends to a final stage which is reached, if ever, at a date that we denote t_{ss}. In the final stage, the carbon budget is binding ($m_t = \bar{m}$), implying that atmospheric emissions are stable ($\dot{m}_t = 0$) and polluting capital is constant at $k_{p,t} = \bar{m}\,\varepsilon/G$.

Before the transition is over, a classical result (see for instance footnote 11 in Goulder and Mathai, 2000) is that the shadow carbon price grows at the interest rate r_t plus the dissipation rate of GHG (Appendix B.2):

$$\forall t,\ m_t < \bar{m} \implies \dot{\tau}_t = \tau_t\,(r_t + \varepsilon) \tag{35}$$

where the endogenous interest rate r_t is defined as the marginal return from clean investments net from depreciation:

$$r_t := \frac{\partial F}{\partial k_c} - \delta \tag{36}$$

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.[6]

In the *laissez-faire* equilibrium, capital was used up to the point where the marginal productivity of polluting capital was equal to its rental rate. This is no longer the case, since the social planner now accounts for the social cost of carbon when they use polluting capital. They must therefore reduce the amount of polluting capital used for production, to increase its marginal productivity:

Lemma 2. *Along the socially-optimal path, the marginal productivity of clean capital equals the implicit rental rate of clean capital:*

$$\frac{\partial F}{\partial k_c} = R^i_{c,t} \tag{37}$$

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

$$\frac{\partial F}{\partial q_p} = R^i_{p,t} + \tau_t\, G \tag{38}$$

---

[6] Rezai and Van der Ploeg (2016) use a more complex climate model and account for fossil reserve depletion, risk aversion, and use a cost-benefit approach, and still find that the optimal carbon price essentially grows exponentially over time before the transition is complete.

Proof. Equation (37) derives from (26) and (33). Equation (38) is obtained by substituting β_t in (28), using (34). □

Another difference with the laissez-faire equilibrium is that the implicit rental rate of polluting capital $R^i_{p,t}$ now differs from that of clean capital $R^i_{c,t}$:

$$R^i_{p,t} = R^i_{c,t} - \frac{1}{\lambda_t}\left((\rho+\delta)\psi_t - \dot{\psi}_t\right) \tag{39}$$

where ψ_t is the Lagrange multiplier associated with the irreversibility constraint. Defining the *legacy cost* ℓ_t as the "annualized" value of the shadow cost of the irreversibility constraint ψ_t; similarly to how the implicit rental rate $R^i_{p,t}$ relates to the value of new polluting capacities ν_t:

$$\ell_t := \frac{1}{\lambda_t}\left((\rho+\delta)\psi_t - \dot{\psi}_t\right) \tag{40}$$

we get:

$$R^i_{p,t} = R^i_{c,t} - \ell_t \tag{41}$$

Starting from the laissez-faire equilibrium, with $R^i_{p,t} = R^i_{c,t}$, the fact that investment is irreversible prevents the planner to adjust the stock of polluting capital instantaneously when the climate objective is imposed. Polluting capital therefore becomes relatively more abundant and the legacy cost imposes a gap between the implicit rental rates of clean and dirty capacities. The legacy cost can be seen as a quantification of the regret that society has because of excessive irreversible investment in polluting capital (e.g. having built a coal power plant before the climate mitigation policy has been announced or before realizing the dangers associated with climate change).

Another manifestation of the irreversibility of polluting investment is that the system can go through two types of phases during the transition to a clean economy:

Lemma 3. *The transition to clean capital can go through two types of phases:*

1. *Phases with some investment in polluting capital, during which there are no legacy costs and the implicit rental rate of polluting capital is equal to the implicit rental rate of clean capital and*

$$\begin{aligned} i_{p,t} &> 0 \\ \ell_t &= 0 \\ R^i_{p,t} &= R^i_{c,t} \end{aligned}$$

2. *Phases with no investment in polluting capital, during which legacy costs are positive and the implicit rental price of polluting capital can be lower than the rental rate of clean capital:*

$$\begin{aligned} & i_{p,t} = 0 \\ & 0 \le \ell_t \le R^i_{c,t} \\ & R^i_{p,t} \le R^i_{c,t} \end{aligned}$$

Proof. The complementary slackness condition 30 implies that if polluting investment is strictly positive, then investment is chosen at the level that equalizes the value of polluting and clean capital:

$$\exists(t_1,t_2)|\forall t \in [t_1,t_2],\ i_{p,t} > 0 \tag{42}$$

$$\implies \forall t \in [t_1,t_2]\ \psi_t = 0 \tag{43}$$

$$\implies \forall t \in [t_1,t_2], \nu_t = \chi_t \tag{44}$$

$$\implies \forall t \in [t_1,t_2],\ R^i_{p,t} = R^i_{c,t} \tag{45}$$

On the other hand if $i_{p,t} = 0$ then $\psi_t > 0$ is possible, and in general $R^i_{p,t} \leq R^i_{c,t}$. (Lemma 6 below exhibits cases when the inequality is strict $R^i_{p,t} < R^i_{c,t}$.) Since $R^i_{p,t} = \beta_t/\lambda_t \geq 0$, $\ell_t = R^i_{c,t} - R^i_{p,t} \leq R^i_{c,t}$. □

As in many Ramsey formulation (e.g., equations 11 to 14 in Arrow and Kurz, 1970), polluting investment enters linearly in the Hamiltonian (B.1). Thus, the FOCs (25–28) do not give a direct rule to choose investment as a function of other state variables at each time step. Lemma 3 shows that we are however left with only two possibilities, which together form what is sometimes called a *bang-singular* solution: phases when investment is null, which are sometimes called *bang* phases, and phases when investment is set indirectly by the value of capital, which are sometimes called *singular* rays.[7]

The next lemma establishes that the transition from the laissez-faire equilibrium to a cleaner economy necessarily starts with a phase with no investment in polluting capital; and finishes with a phase with some investment in polluting capital – such that the legacy cost are only temporary, because in the long term, excess polluting capital has depreciated to a sustainable level.

Lemma 4. *The transition necessarily starts with a phase featuring no polluting investment, and ends with a phase featuring some polluting investment and equal rental rates of clean and polluting capital.*

Proof. Appendix B.3. □

Polluting capital in our model behaves similarly to the capital in the Ramsey model with irreversible investment studied by Arrow and Kurz (1970). A new finding in our paper is that the irreversibility constraint on polluting capital always becomes binding when, from the laissez-faire equilibrium, an unanticipated constraint on GHG concentration is suddenly imposed.

Lemma 3 also means that in the social optimum, the maximum possible value for the legacy cost ℓ_t is the marginal productivity of clean capital $\frac{\partial F}{\partial k_c}(= R^i_{c,t})$: at worst, the social planner regrets not to have invested in clean instead of

---

[7] Note that Lemma 3 was derived from a study of the Lagrange multiplier associated with irreversibility constraint and the complementary slackness condition in the spirit of Karush–Kuhn–Tucker's nonlinear programming theory. Arrow and Kurz (1970) use another technique, equivalent in this case, in the spirit of singular control theory.

polluting capital before t_0. In that case, the implicit rental rate of polluting capital falls down to zero, reflecting that polluting capital is over-abundant and should be underused:

Lemma 5. *Polluting capital is underutilized if and only the carbon price is greater than the marginal productivity of installed polluting capital divided by its carbon intensity,*

$$\tau_t \, G > \frac{\partial F(k_{p,t}, k_{c,t})}{\partial k_p} \implies \begin{cases} q_{p,t} < k_{p,t} \\ \ell_t = R^i_{c,t} \\ R^i_{p,t} = 0 \\ \frac{\partial F(q_{p,t}, k_{c,t})}{\partial q_p} = \tau_t \, G \end{cases} \tag{46}$$

$$q_{p,t} = k_{p,t} \implies \tau_t \, G \leq \frac{\partial F(k_{p,t}, k_{c,t})}{\partial k_p} \tag{47}$$

Proof. (38) implies that the implicit rental rate of polluting capital $R^i_{p,t}$ is the difference between the marginal productivity of polluting capital and social cost of carbon. As the implicit rental rate of polluting capital $R^i_{p,t}$ is equal to the positive Lagrange multiplier associated to the capacity constraint β_t ((27) and (34)), and given the complementary slackness condition (31), it is easy to show that when the carbon price is greater than the marginal productivity of installed polluting capital the implicit rental rate of polluting capital is nill and capital is underutilized. Reciprocally, $q_{p,t} = k_{p,t} \implies \beta_t = R^i_{p,t} \geq 0 \implies \tau_t \, G \leq \frac{\partial F(q_{p,t}, k_{c,t})}{\partial q_p} = \tau_t \, G \leq \frac{\partial F(k_{p,t}, k_{c,t})}{\partial k_p}$ □

Lemma 6 means that stopping to use some of the polluting capital that was constructed before the climate policy is enacted can be part of the optimal strategy to reduce the cost of the transition to clean capital. Since in our framework all polluting capital is in aggregate, this lemma can be interpreted as an underutilization of the whole stock of polluting capital, or the premature retirement of a fraction of the available capacities. (In practice capacities are not homogeneous, and the most polluting units of capital, for instance the oldest coal power plants, could be decommissioned first). The next subsection shows that the drop in the implicit rental rate of capital translates into stranded assets for the owners of polluting capital when the social optimum is enforced with a carbon price.

Note that polluting capacities are not under-used during all of the transition, and in particular:

Lemma 6. *Polluting capacities are fully used during phases when there is investment in polluting capital*

Proof. $i_{p,t} > 0 \implies R^i_{p,t} = R^i_{c,t} \implies \beta_t > 0 \implies q_{p,t} = k_{p,t}$ □

Underutilization of polluting capital can happen at the beginning of transition, depending on the GHG concentration ceiling $\bar{m}$, on the initial stock of

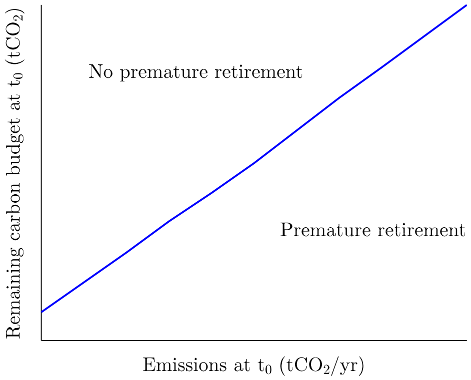

Figure 1: Under-utilization of polluting capital as a function of initial emissions and the ceiling. Depending on initial emissions (given a $k_{b,0}$, these depend directly on G) and on the concentration ceiling ($\bar{m}$), polluting capital is underutilized or not in numerical simulations of the least-cost transition.

polluting capital k_{p,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.

This article is illustrated with simulations from a numerical implementation of the social planer problems (Appendix E provides the code used to run the model). We select simulations and figures to illustrate some insights from the analytical resolution. Figure 1 illustrates how given a set of functions and parameters, the underutilization of polluting capital happens if initial endowment of polluting capital results in high initial emissions (right end of the x-axis) and/or if the atmospheric carbon ceiling is stringent (lower part of the y-axis). The figure was obtained solving the model numerically in various simulations, varying G and $\bar{m}$ while maintaining all the other parameters and functional forms constants.

We summarize the findings about the socially-optimal transition to clean capital in the following proposition, illustrated by Figure 2:

Proposition 1. *The optimal transition from the laissez-faire equilibrium to the efficient final stage goes through phases of two kinds.*

1. *In the first type of phase, the irreversibility of investment translates into a gap between the rental rate of polluting and clean capital. As a result, no investment goes to polluting capital during this phase. At worst, the implicit rental rate of polluting capital can drop to zero, and existing polluting capacities can be underused.*
2. *In the second type of phase, the rental rates of clean and polluting capital are equal and polluting capital is utilized in full.*

The optimal transition necessarily starts with a phase with no investment in

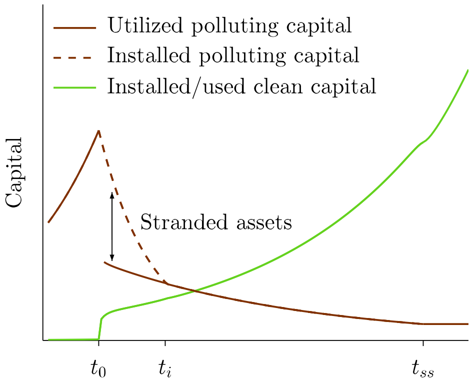

Figure 2: Installed polluting and clean capital, and utilized polluting capital in the least-cost transition to clean capital Before t_0, the economy is on the *laissez-faire* equilibrium, during which the stock of clean capital is small but not null. At t_0 the carbon price is implemented, investment in polluting capital stops, and polluting capital depreciates until t_i ($\forall t \in (t_0, t_i),\ i_p = 0$). During this period, a portion of polluting capital may be underutilized or retired prematurely ($q_{p,t} < k_{p,t}$). Here, the final stage is reached at t_{ss}.

polluting capital, potentially featuring premature retirement, and ends with a phase with full utilization of polluting capacities.

550 Proof. Lemmas 4 and 6. □

The concepts of premature retirement and legacy costs can also be used to decompose the social cost of carbon. From equations (37), (38) and (41), we get:

$$\tau_t = \frac{\frac{\partial F}{\partial q_p}(q_{p,t}, k_{c,t}) - \frac{\partial F}{\partial k_c}(q_{p,t}, k_{c,t}) + \ell_t}{G}$$

Resting and adding $\frac{\partial F}{\partial q_p}(k_{p,t}, k_{c,t})$, yields:

$$\begin{aligned} \implies \underbrace{\tau_t}_{\text{Marginal abatement cost}} &= \underbrace{\frac{\frac{\partial F}{\partial q_p}(q_{p,t}, k_{c,t}) - \frac{\partial F}{\partial q_p}(k_{p,t}, k_{c,t})}{G}}_{\text{Premature-retirement cost}} \\ &+ \underbrace{\frac{\frac{\partial F}{\partial q_p}(k_{p,t}, k_{c,t}) - \frac{\partial F}{\partial k_c}(q_{p,t}, k_{c,t})}{G}}_{\text{Technical cost}} \\ &+ \underbrace{\frac{\ell_t}{G}}_{\text{Legacy cost}} \end{aligned} \tag{48}$$

555 The social cost of carbon τ_t decomposes as a cost of premature retirement (the difference between the marginal productivity of polluting capital were all

capacities utilized and the marginal productivity of polluting capital given that all polluting capacities are not utilized), a technical cost (for instance existing renewable power plants are less economically efficient than existing coal power plants if all coal plants were used), and the legacy cost.

In this section, we thus 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. In the next section, we show that a carbon price can decentralize the social optimum, and that in that case legacy costs directly impact the current owners of polluting capital, in the form of stranded assets. Then, we turn to a social planner program where premature retirement is to be avoided, and show that who pays the legacy costs (and whether stranded assets are avoided) in that case depends on what policy instruments are used to decentralize that program.

4.2. Decentralization with a Carbon Tax

Unsurprisingly, the social planner can trigger the same outcome as in the social optimum in a decentralized economy by imposing a price on carbon emissions. Starting from the decentralized model exposed in Section 3.1, the firm's flow of profit (10) is modified to:

$$\Pi_t = F(A_t, q_{p,t}, k_{c,t}) - R_{c,t} \cdot k_{c,t} - R_{p,t} \cdot q_{p,t} - \tau_t \, G \, q_{p,t} \tag{49}$$

Where τ_t is a carbon tax schedule numerically equal to the optimal shadow carbon price in the social planner's program (35).

The FOCs for the producer become:

$$\frac{\partial \Pi}{\partial k_c} = 0 \implies \quad R_{c,t} = \frac{\partial F}{\partial k_c}(q_{p,t}, k_{c,t}) \tag{50}$$

$$\frac{\partial \Pi}{\partial q_p} = 0 \implies \quad R_{p,t} = \frac{\partial F}{\partial q_p}(q_{p,t}, k_{c,t}) - \tau_t \, G \tag{51}$$

The problem for the household is essentially unchanged. First order conditions for the household are the same as in the laissez-faire (13), (14), (15). Combining them with the FOCs for the producer leads to the same set of equations than the FOCs from the planner's program in the previous section — with the implicit rental costs of capital ($R^i_{p,t}$, $R^i_{c,t}$) replaced by the actual rental costs of capital ($R_{p,t}$, $R_{c,t}$), and the social cost of carbon replaced by the actual price of carbon. This means that the carbon price leads to the socially-optimal investment and production decisions.

Applied to the decentralized equilibrium, results from the previous section mean that when the government implements a carbon price, the actual rental rate of polluting capacities is affected by the legacy costs:

$$R_{p,t} = R_{c,t} - \ell_t \tag{52}$$

In that sense, the legacy costs ℓ_t are paid by the owners of stranded assets: an unanticipated carbon price creates a sudden gap between revenues from clean and dirty capacities. At worst, in case of premature retirement of polluting capacities, the rental rate of polluting capacities can be reduced to zero (lemma 6).

The socially-optimal carbon price may thus turn out to be politically difficult to implement, as it imposes immediate and concentrated costs on a few players, the owners of polluting capital, who could organise and oppose the reform (Olson, 1977; Trebilcock, 2014); while its benefits, avoided climate change, are diffuse over all actors and over time, which tends to reduce mobilization to defend the reform.

In the next section, we solve for a constrained equilibrium where premature retirement is to be avoided for political reasons. This increases the total economic cost of the transition, but who pays that cost depends on which specific instruments the government uses to enforce the second-best transition.

5. Avoiding Premature retirement and Stranded Assets

5.1. Social Planner Program

Here, we solve a new social planner program, identical to the first best optimum, but with the additional political constraint that polluting capital should not be underused:

$$\max_{c,i,q} \int_0^\infty e^{-\rho t} \cdot u(c_t)\, dt \tag{53}$$

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

In this problem, we have left two constraints for analytical purposes: the physical one, that capacity cannot be overused, and the political choice that underutilization should not occur. The latter is expressed as an inequality which is binding only when there would otherwise be premature retirement, rather than an equality, also for analytical tractability.

First-order conditions are available at Appendix C. The following complementary slackness conditions play a key role:

$$\forall t,\ \beta_t \geq 0 \text{ and } \beta_t \cdot (k_{p,t} - q_{p,t}) = 0 \tag{54}$$

$$\forall t,\ \alpha_t \geq 0 \text{ and } \alpha_t \cdot (k_{p,t} - q_{p,t}) = 0 \tag{55}$$

The system tends towards the same final stage as in the first-best case. During the transition to that final stage, the social cost of carbon still grows at the interest rate net of carbon dissipation rate.

However, the no-underutilization constraint changes the relationship between the shadow cost of the capacity constraints and the value of new capacities captured by (27) in the first-best equilibrium. The new relationship reads:

$$\beta_t - \alpha_t = ((\delta + \rho)\nu_t - \dot{\nu}_t) \tag{56}$$

So that defining the implicit rental rates of capital as before (33, 34) now yields $\lambda_t R^i_{p,t} = \beta_t - \alpha_t$. The implicit rental cost of polluting capital depends on both the physical constraint that capacities cannot be *over*used and the political constraint that they shall not be *under*used. The implicit rental cost of clean capital is as in the first best. Equations (37) and (38) are unchanged: the marginal productivity of polluting capital is still the sum of the implicit rental price of capital and the social cost of carbon, thus:

$$\frac{\beta_t - \alpha_t}{\lambda_t} = \frac{\partial F}{\partial q_p} = R^i_{p,t} + \tau_t \cdot G \tag{57}$$

However, the implicit rental rate of polluting capital $R^i_{p,t}$ can now be negative. Since both β_t and α_t are positive by construction, and $\lambda_t = u'(c_t) > 0$,

$$\alpha = 0 \implies \frac{\beta}{\lambda_t} = \frac{\partial F}{\partial q_p} - \tau_t \cdot G \geq 0 \implies R^i_{p,t} \geq 0 \tag{58}$$

$$\beta_t = 0 \implies \frac{\alpha}{\lambda_t} = \tau_t \cdot G - \frac{\partial F}{\partial q_p} \geq 0 \implies R^i_{p,t} \leq 0 \tag{59}$$

These equations reflect the following. At each point in time t, three cases captured by the complementary slackness conditions are possible, depending on how much polluting capital the social planner would use in the absence of political and physical constraints: (i) the planner would use exactly the amount of polluting capital that is currently available; in this case we say that neither constraint is *binding*; (ii) the planner would use more polluting capital than what is currently available; in this case the physical constraint is binding; (iii) the planner would use less polluting capital than what is currently available; in this case the political constraint is binding.

In particular, if the capacity constraint is not binding, then $\beta_t = 0$. This implies that the social cost of carbon is higher than the marginal productivity of polluting capital expressed in dollars per ton, and that the constraint that polluting assets shall not be underused is binding ($\alpha_t \geq 0$). In that case, the implicit rental rate of polluting capital is negative. The next subsections show that depending on what policy instruments the government uses to enforce the constrained transition, α may or may not materialize as a subsidy to the utilization of polluting capital.

On the other hand, when the political no-underutilization constraint is not binding, then $\alpha_t = 0$ and, as expected, the implicit rental rate of polluting capital behaves as in the first-best equilibrium ($\beta = R^i_{p,t} > 0$).

Identically to what happens in the first-best pathway, the marginal productivities are differentiated by legacy costs and a term proportional to the social cost of carbon:

$$\frac{\partial F}{\partial q_p} = \frac{\partial F}{\partial k_c} - \ell_t + \tau_t \, G \tag{60}$$

where legacy costs ℓ_t are defined as previously (40), yielding $R^i_{p,t} = R^i_{c,t} - \ell_t$. In the constrained transition, decomposing the social cost of carbon τ_t as before ((48)) now yields:

$$\Longrightarrow \quad \underbrace{\tau_t}_{\text{Marginal abatement cost}} = \underbrace{\frac{\frac{\partial F}{\partial q_p}(k_{p,t}) - \frac{\partial F}{\partial k_c}}{G}}_{\text{Technical cost}} + \underbrace{\frac{\ell_t}{G}}_{\text{Legacy cost}} \tag{61}$$

The cost of premature retirement is now zero by construction. (But of course, that does not lead to a lower abatement cost). On the other hand, the legacy cost is no longer bounded by $R^i_{c,t}$ as in lemma 4, because $R^i_{p,t}$ can now be negative. In particular, equation (60) shows that at the beginning of the transition, the legacy cost equals the carbon price.

With the first-best carbon price, the maximum regret linked to excess past installation of polluting capital was the opportunity cost of not having invested in clean capital. Here, 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 legacy cost can be as high as the cost of the carbon emissions generated by the polluting capital, which is higher than in the first-best transition. Refusing to strand assets thus increases regret from past investment in those assets, as their utilization make the climate target more difficult to achieve. The next subsections show that who pays the legacy costs depends on policy design.

As in the first-best case, the constrained transition starts with a phase with no investment in polluting capacity. During this phase, since growth happens from clean capital accumulation, a gap between the productivity of clean and polluting capital can appear, and grow over time $\frac{\partial F}{\partial q_p} > \frac{\partial F}{\partial k_c}$, giving space for legacy costs to decrease over time (60). Before the carbon budget is depleted, there may also be phases where investment in polluting capital is strictly positive (capacity variation net of depreciation can remain negative during those phases). Finally, the system reaches the same final stage as previously.

Since capacities are not underused, short-term output may be higher in the constrained transition than in the first-best strategy. Figure 3 illustrates this result. Analytically, the effect on consumption is ambiguous because it involves the offsetting impacts from an income effect (short-term output is higher) and two substitution effect (investment in clean capital is cheaper, which tends to decrease short-term investment and thus increase consumption, and investment in polluting capital is more expensive, which tends to increase short-term consumption).

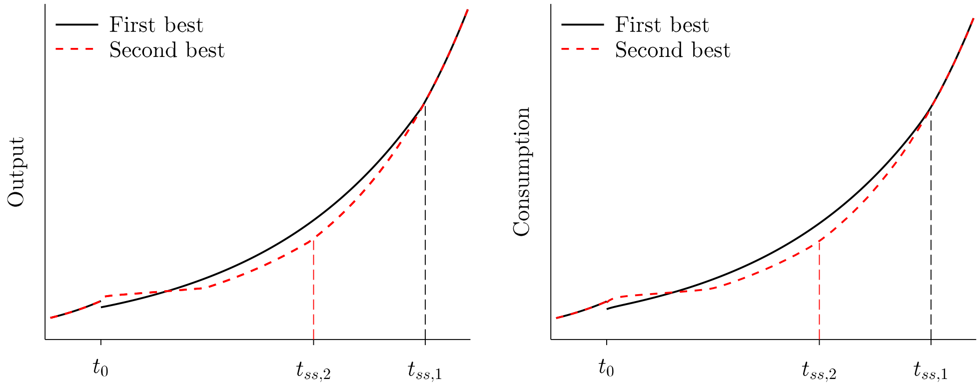

Figure 3: Output in the two simulations. The figure shows output y in the first-best and the constrained transitions. In the short run, output is lower in the first-best case because of premature retirement of polluting capacities. On the particular example in the right, short-term consumption c is higher in the second-best case because of a higher output y. $t_{ss,1}$ and $t_{ss,2}$ are the dates when the final stage is reached in the first best and second best transitions respectively. In the figure the final stage is reached sooner in the second-best case ($t_{ss,2} < t_{ss,1}$).

Figure 4 compares the shadow cost of carbon in an optimal and a constrained transition using the same calibration (Appendix E). Avoiding premature retirement generates a higher social cost of carbon than the first-best carbon price. However the dynamics of capital accumulation mean that the social cost of carbon at each point in time does not translate into immediate consumption losses at the same point in time (Vogt-Schilb et al., 2018). In this case, while the constrained transition sets a higher shadow cost of carbon at each time t (figure 4), they lead to higher output and possibly consumption over the short-run (figure 3).

Results from this section are summarized in the following proposition:

Proposition 2. *The constrained transition leads to the same final stage as the optimal transition, the constrained pathway thus differs only temporarily from the first-best pathway.*

The constrained transition imposes a higher shadow cost of carbon, and initially higher legacy costs than the optimal transition.

Compared to the optimal transition, the constrained transition smooths social costs: it decreases effort in the short-run, leaves them unchanged in the long-run (as the final stage remains unchanged), and thus increases effort in the medium-run.

In the following section, we show that feebates, mandates, and phased-in carbon prices can all decentralize the constrained transition. But feebates and mandates protect revenues for the owners of existing polluting capital, while the second-best phased-in carbon price does not.

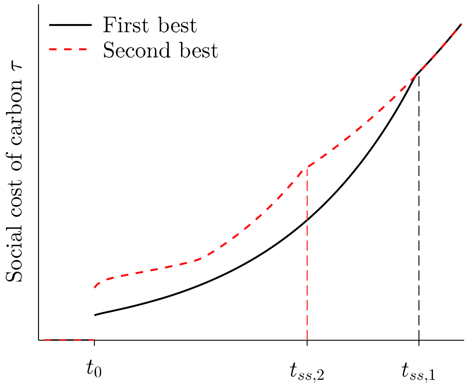

Figure 4: Social cost of carbon in the two simulations. The shadow price of emissions τ is higher in the constrained transition. The dates $t_{ss,1}$ and $t_{ss,2}$ denote the moment when the first best transition and the second best transition, respectively, reach the final stage.

5.2. Decentralization of the Second-best Equilibrium Combining a Carbon Price and a Subsidy on Polluting Production, or a Phased-in Carbon Price

One way to decentralize the constrained optimum is to simply translate the social cost of carbon τ and the shadow subsidy against premature retirement α into an actual carbon tax and an actual subsidy on polluting production.

In that case, the program of the producer (10) becomes:

$$\max_q \pi_t = F(A_t, q_{p,t}, k_{c,t}) - R_{c,t}k_{c,t} - (R_{p,t} + \tau_t G - \alpha_t)q_{p,t} \tag{62}$$

This leads to the FOCs for the producer:

$$R_{c,t} = \frac{\partial F}{\partial k_c} \tag{63}$$

$$R_{p,t} = \frac{\partial F}{\partial q_p} - \tau_t G + \alpha_t \tag{64}$$

And the household problem is essentially unchanged.

If the government sets the value of the carbon tax schedule and the subsidy equal to the optimum values of the respective Lagrange multipliers from the previous section, the set of FOCs is identical to the one from the constrained transition, where implicit rental rates have been replaced by actual rental rates, the social cost of carbon is replaced by the carbon price, and the shadow value of the no-underutilization constraint is replaced by a subsidy on dirty capital usage, leading to the same transition to a clean economy.

Since the FOCs of the producer only depend on $\tau_t G$ and α_t via $(\tau_t G - \alpha_t)$, the government can also decentralize the constrained transition using a single instrument: the phased-in tax scheduled $\tilde{\tau}_t = \tau_t G - \alpha_t$. At the beginning of the transition, $\alpha_t > 0$ reduces the phased-in carbon price below its first-best schedule. Note that the phased-in carbon price increases over time for two reasons: first, efficiency conditions mean that the carbon price should increase, basically at the interest rate, as long as the transition to a clean economy is not complete (35). And second, to avoid premature retirement, the actual carbon price would start at a lower-than-efficient value, and catch-up with the optimal value to give gents time to adjust to the new prices.

Premature retirement can thus be avoided using a phased-in carbon price, or, equivalently, the combination of a carbon price and a subsidy for production from polluting capacity. Both instruments achieve this by imposing an effective price of carbon, $(\tau_t - \alpha_t/G)$, which is lower than the social cost of carbon τ_t.

It does not follow, however, that these instruments are harmless for the owners of polluting capacities. Indeed, the previous subsection shows that the highest possible value for the second-best subsidy α_t is $\tau_t G - \frac{\partial F}{\partial q_p}$. But in that case, the subsidy covers exactly the gap between marginal productivity of polluting capital and the carbon price, implying that market price for renting polluting capacities is still zero $R_{p,t} = 0$ during this phase (64). Even if it avoids early retirement, the second-best phased-in carbon price does not avoid stranded assets.

This results nuances the claim by Williams (2012) that phasing-in a carbon price is a way of dealing with distributional impacts of climate policies. Our model suggests that while a phased-in carbon tax can avoid premature retirement and an economy-wide drop in production when it is announced, it is set at the level where owners of polluting capacity are indifferent between renting out their capacities or scrapping them. Therefore, the second-best phased-in carbon price does not automatically protect the revenues of the owners of polluting capacities, that is it can create stranded assets.

To protect the revenues from polluting capital, the government could use an even more gradual carbon price, but that would necessarily result in higher social costs than using the second-best schedule designed to minimize costs while avoiding premature retirement. (The government could also offset the losses of the owners of polluting capital with ex-post lump-sum transfers.) In contrast, the next section shows that if government use instruments that regulate new investment decisions instead of production decisions, such as feebate programs or standards on new equipment, then the same transition to a clean economy can be enforced while protecting ex-ante the revenues of the owners of polluting capital.

We summarize the findings of this section in the following lemma:

Proposition 3. *A phased-in carbon tax is equivalent to a carbon tax complemented with a temporary subsidy on polluting capacity. Both instruments allow decentralization of the second-best transition to clean capital where premature retirement is to be avoided. Both instruments can lead to stranded assets, however, and at worst the revenues from existing polluting capital can drop down to zero.*

5.3. Decentralization with Feebate or Mandates on New Investment

Current climate mitigation policies are not limited to carbon prices; many governments rely instead on instruments such as energy efficiency standards, direct public investment in "green" sectors such as public transport, and fiscal incentives for green investment such as feebates, which impose additional fees on polluting capital and rebates for clean capital (IEA, 2016). These instruments redirect investment towards clean capital but have no effect on the use of existing capital.

MANDATES. One way to regulate investment is by imposing a moratorium on polluting investment and mandating investment in clean capacity. We call these instruments *mandates* on new investment.

Mandates may seem extreme in our model, but similar instruments are actually used by policy-makers and discussed in the field of climate policy. For instance, many jurisdictions have effectively banned incandescent light bulbs or personal vehicles with very low energy efficiency (IEA, 2016). Other jurisdictions have banned nuclear power plants. Bertram et al. (2015) and Pfeiffer et al. (2016) propose to ban the construction of new standard coal and gas power plants, and to mandate new power plants to be renewable power, nuclear, or fossil fuel plants equipped with carbon capture and storage.

Note that while our model only represents "clean" and "dirty" capital, the actual implementation of mandates is not necessarily a pure command-and-control policy, because several types of clean capital may be available to comply (Azar and Sandén, 2011). For instance, a renewable electricity mandate still lets the market chose amongst a range of wind or solar power technologies. Some standards work similarly to mandates: for instance minimum energy efficiency on lighting can in effect ban incandescent light bulbs and mandate LEDs instead.

With mandates, the household problem in our model becomes:

$$\max_{c,i,q} \int_0^\infty e^{-\rho t} \cdot u(c_t)\, dt \tag{65}$$

$$\begin{aligned}
\text{subject to } & R_{c,t} \cdot k_{c,t} + R_{p,t} \cdot q_{p,t} = c_t + i_{p,t} + i_{c,t} && (\lambda_t) \\
& \dot{k}_{p,t} = i_{p,t} - \delta k_{p,t} && (\nu_t) \\
& \dot{k}_{c,t} = i_{c,t} - \delta k_{c,t} && (\chi_t) \\
& i_{p,t} \geq 0 && (\psi_t) \\
& q_{p,t} \leq k_{p,t} && (\beta_t) \\
& i_{p,t} \leq s_{p,t} && (\sigma_{p,t}) \\
& i_{c,t} \geq s_{c,t} && (\sigma_{c,t})
\end{aligned}$$

The mandates $s_{p,t}, s_{c,t}$ can be set to equal polluting investments found in section 5.1. In this model, $s_{p,t} = 0$ until polluting capacities have depreciated to a level compatible with the carbon ceiling.

Notice that two instruments may be needed here: a moratorium on polluting investment alone imposes a shadow price on investment, and can thus result in the household consuming too much and saving too little, compared to the second-best constrained transition. The mandate on clean investment compensates that. (In the Appendix C.2 we show that performance standards formulated on averages, such as the CAFE standards in the US, in contrast, cannot always decentralize the second-best transition).

With mandates, the firms problem is the same as in the laissez-faire, now implying that firms are always willing to pay a strictly positive rent for both clean and dirty capacities: equations (19) and (20) hold.

In particular, the comparison between rental rates for clean and dirty capital does not involve the legacy cost any more. The rental rate on polluting capital therefore does not drop when the policy is implemented, and remains strictly positive during the transition. Strictly positive rental rates also imply that with mandates on new investment, the household always rents out all the available capital: there is no premature retirement.

In fact, revenues from polluting capacities may be boosted by these instruments. Consider how the gap between clean and polluting capital evolves over time:

$$\dot{R}_{p,t} - \dot{R}_{c,t} = (\dot{k}_{c,t} - \dot{q}_{p,t})\frac{\partial^2 F}{\partial k_c \partial q_p} - \dot{q}_{p,t}\frac{\partial^2 F}{\partial {q_p}^2} + \dot{k}_{c,t}\frac{\partial^2 F}{\partial {k_c}^2} \tag{66}$$

A sufficient condition for the above to be a sum of positive terms is that the stock of dirty capital decreases over time ($\dot{q}_{p,t} < 0$) and the stock of clean capital increases over time ($\dot{k}_{c,t} > 0$). By limiting the supply of new dirty capacities (but not tying a cost to the usage of those capacities, like a carbon tax would) and boosting the supply of new clean capacities, the government is creating windfall profits and favouring the vested interests that already own polluting capacities.

Since mandates constrain investment decisions and thus capital stocks, since the household rents out all of those stocks, and since consumption equals total production net of investment, we have shown that:

Lemma 7. *In our model, well-designed mandates can decentralize the constrained transition while avoiding stranded assets.*

FEEBATE. Unsurprisingly, the same transition can be obtained using price instruments, for instance a so-called feebate program. A feebate is the combination of a subsidy (or rebate) $\theta_{c,t}$ on investment in clean capacity and a tax (or fee) $\theta_{p,t}$ on investment in polluting capacity.[8]

With feebates, the household problem becomes:

$$\max_{c,i,q} \int_0^\infty e^{-\rho t} \cdot u(c_t)\, dt \tag{67}$$

subject to $B_t + R_{c,t} \cdot k_{c,t} + R_{p,t} \cdot q_{p,t} - c_t - i_{p,t}(1+\theta_{p,t}) - i_{c,t}(1-\theta_{c,t}) = 0 \quad (\lambda_t)$

---

[8] Our model does not capture all important factors in the choice between mandates and feebates. For instance mandates fail to impose the same marginal abatement cost to heterogeneous producers and therefore cannot be efficient (e.g., Anderson et al., 2011).

$$\dot{k}_{p,t} = i_{p,t} - \delta k_{p,t} \qquad (\nu_t)$$

$$\dot{k}_{c,t} = i_{c,t} - \delta k_{c,t} \qquad (\chi_t)$$

$$i_{p,t} \geq 0 \qquad (\psi_t)$$

$$q_{p,t} \leq k_{p,t} \qquad (\beta_t)$$

Where B_t is the net budgetary impact of the feebate scheme, taken as exogenous by the representative household. As before, the scheme is not necessarily revenue neutral, and we assume that the net revenue (net cost) from the scheme is recycled (financed with taxes) in a lump-sum way that has no impact on production, investment and consumption decisions at the margin.

To decentralize the constrained social optimum, the government simply needs to set the feebate $(\theta_{p,t}, \theta_{c,t})$ such that the values of clean and polluting investment are the same as in the case of mandates (Appendix C.1).

Moreover, the firms problem remains unchanged. With a feebate, rents are positive, and there are no stranded assets; the rental rate of polluting capacities does not drop when the policy is implemented, and can even increase relative to the rental rate of clean capacities:

Proposition 4. *The constrained transition to clean capital where premature retirement is to be avoided can be decentralized with mandates on new investment, or with feebates on new investment. Such second-best instruments do not directly produce stranded assets.*

PROOF. Appendix C.1 provides more details.

In summary, section 5 has shown that at least four different policy instruments can decentralize the constrained transition where early retirement of polluting capital is to be avoided: i) a carbon price and a subsidy on polluting production, ii) a phased-in carbon price, iii) mandates on new clean and polluting investments, and iv) a tax/subsidy scheme - a feebate - on new investments in polluting and clean capital. All these instruments can be designed so they lead to the same set of consumption, production and investment decisions. They thus lead to the same social cost, and the social planner of our model is indifferent between them.

Different instruments however affect the producer and the investor-consumer of our model differently. The phased-in carbon price and the combination of a carbon tax with a subsidy (options i and ii) both decrease the rental rate of polluting capital, possibly making it drop to zero. In contrast, the mandates (option iii) and feebates (option iv) always lead to strictly positive rents for the owners of polluting capital – and can even increase the rent of polluting capital relative to clean capital.

These four second-best instruments can decentralize the constrained transition at the same total cost, but not necessarily with the same distribution of this cost.[9] With net fiscal revenues from tax instruments given back lump-sum

---

[9] As mentioned in footnote 2, the fiscal impact of different policy instruments, while im-

to the consumer-investor, any instrument choice that results in extra costs for the producer in our model is an extra gain for the consumer-investor. In our model, the producer prefers feebates and mandates (that have a direct impact on investment decisions) over the phased in carbon tax (that has a direct impact on production decisions). The consumer-investor has the opposite preference.

A more sophisticated ranking of those instruments is not in the scope of this paper based on a simple analytical model. Nonetheless, our results provide insights on the possible social acceptability of different instruments, in particular on the fact that some instruments directly lead to stranded assets, while others do not.

Results exposed in this section also extend previous research (e.g., Fischer and Newell, 2008; Holland et al., 2009) that uses static models and finds that performance standards and feebate schemes act as the combination of a carbon tax and a production subsidy. With our dynamic model, we have clarified that while this shadow subsidy protects production and revenues from pre-existing polluting capital, it does not provide incentive to invest in additional polluting capital. The effect of the shadow subsidy is thus only temporary since once the level of polluting capital has decreased to a sustainable path, all instruments are equivalent to a simple carbon tax. And, perhaps more surprisingly, the incidence of a carbon price plus a temporary subsidy on polluting production differs from the incidence of mandates or feebates on investment decisions (Since this result required to model separately investors and producers, it was not found in the above-mentioned papers.)

6. Committed Emissions, Instrument Choice, and Carbon Lock-in

Since they maintain a full utilization of polluting capital in the short term, mandates, feebates, and the second-best phased-in carbon price result in higher short-term emissions than the carbon tax (figure 5). These instruments may thus not be sufficient to reach stringent climate objectives if past accumulation of polluting capital is substantial.

Figure 6 offers a visualization of this issue. At low polluting capital stocks (thus low emissions), a carbon tax does not lead to underutilization of polluting capital. In this case, the first-best carbon price leads to the exact same pathway as second-best mandates or feebates (and the phased-in carbon price is simply equal to the optimal carbon price). This is a situation of flexibility in which a government can enforce the optimal transition to clean capital using any of the instruments discussed in this paper.

But as long as climate policies are absent or too lax, the economy accumulates polluting capital, making GHG emissions grow and reducing the remaining carbon budget for a given climate target (the *laissez-faire growth* arrow).

At one point, the threshold when the marginal productivity of polluting capital is lower than the optimal carbon price is crossed (lemma 6), meaning

---

portant in actuality, is out of scope of this paper.

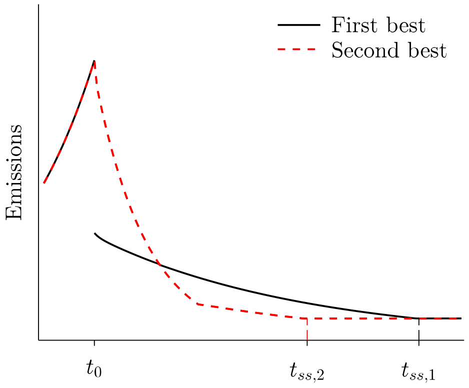

Figure 5: GHG emissions in the two cases. The first-best carbon prices induces decommission of polluting capital and can thus reduce carbon emissions faster than second-best alternative instruments.

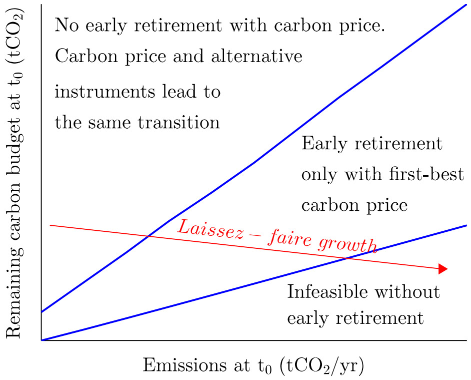

Figure 6: Under-utilization of polluting capital and feasibility of the climate target avoiding premature retirement assets as a function of initial emissions.

that polluting capital should be underutilized and output reduced along the optimal pathway. From there, a carbon price may become even more difficult to implement because of political-economy constraints. But the alternative option of using feebates, mandates or phase-in is still available to reach the same carbon budget without immediate drop in income.

There is thus a window of opportunity, during which alternative policy instruments may induce a smooth and maybe politically-easier transition to a low-carbon economy. If this occasion is missed (bottom and right hand side, figure 6), it becomes impossible to reach the climate target without underutilization of polluting capital and the alternative instruments are not an option anymore (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 involves stranded assets and thus has a significant short-term cost, perhaps 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 figure 6 is expressed analytically in Appendix D and can be approximated by:

$$\bar{m} = m_0 + \frac{G\,k_{p,0}}{\delta}$$

According to Davis et al. (2010), the level of existing polluting infrastructure in 2010 was still low enough to achieve the 2°C target without underutilizing polluting capital. 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 about 1.3°C. While they 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 figure 6 — several studies based on integrated assessment models investigate this question. Rogelj et al. (2013) and Johnson et al. (2015) both find that, in most 2°C scenarios coal power plants are decommissioned before the end of their lifetime, suggesting that the global economy is in the middle zone in figure 6.

In other words, empirical evidence from a few years ago suggests the optimal pathway to a stabilization of the climate at 2°C involves decommissioning existing capital, but that we could still be able to get there by only reducing the carbon content of new capital — in a recent numerical simulation, Bertram et al. (2015) find that a mix between low carbon prices and technology mandates (in particular a moratorium on coal power plants and a minimum requirement for clean power investment) could indeed deliver the 2°C while substantially limiting premature retirement. For some higher temperature target, feebates or mandates and carbon prices are equivalent; while lower temperature targets, such as a 1.5°C target, may now be out of reach if stranded assets are to be avoided — taking into account that since the study by Davis et al. (2010), investment in polluting capital has kept growing and has added to committed

GHG emissions (Davis and Socolow, 2014; Pfeiffer et al., 2018).

7. Conclusion

The present analysis should be interpreted cautiously, as we only explored a few aspects of the transition to clean capital. In particular, our model ignores uncertainty, limited foresight from investors, and limited ability to commit from governments, which can all have important consequences on the comparison between carbon prices, phased-in carbon prices, mandates, feebates and standards regulating present-day investment. One possibility for further research is to integrate and quantify the effect of these elements in a single framework.

Despite these limitations, our results highlight that policy makers face a trade-off between a higher intertemporal efficiency with the optimal carbon price and fewer stranded assets (and perhaps less political costs) with second-best instruments, such as carefully-designed mandates or efficiency standards for new power plants, buildings and appliances, moratoriums on the most carbon-intensive types of capital, *feebate* programs that tax energy-inefficient equipment and subsidize energy-efficient equipment, subsidized loans and tax breaks for energy efficiency investment, or to a lesser extent a phased-in carbon price.

All these instruments are similar in that they redirect private investment away from polluting capital and toward clean capital without providing incentive to drive less or shut down existing coal power plants, preserving revenues from existing polluting capital, and without producing premature retirement of polluting capital. And as they transform progressively the production system, these instruments might prepare the economy and the public to easier implementation of carbon prices in the medium term.

Finally, our results are theoretical findings. They show that in principle government could use alternative instruments to avoid stranded assets and premature retirement in the transition to clean capital. Further research could assess numerically the trade-offs between avoiding private stranded assets and minimizing social costs in any specific market.

To conclude, the analysis carried here may also be relevant for studying other public economy issues. In essence, we propose a parsimonious model able to analyse structural change triggered by policy changes, its impact on vested interests, and policies to manage the transition. Similar models could be used to study policy reform in other topics, such as deregulation of prices in developing markets, trade liberalization, or the advent of robots in labour-intensive sectors.

References

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

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.

André, F. J., Smulders, S., 2014. Fueling growth when oil peaks: Directed technological change and the limits to efficiency. European Economic Review 69, 18–39.

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

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

Asheim, G. B., 2012. A Distributional Argument for Supply-Side Climate Policies. Environmental and Resource Economics 56 (2), 239–254.

Austin, D., Dinan, T., 2005. Clearing the air: The costs and consequences of higher CAFE standards and increased gasoline taxes. Journal of Environmental Economics and Management 50 (3), 562–582.

Azar, C., Sandén, B. A., 2011. The elusive quest for technology-neutral policies. Environmental Innovation and Societal Transitions 1, 135–139.

Barro, R. J., Sala-i-Martin, X. I., 2003. Growth Models with Consumer Optimization (the Ramsey Model). In: Economic Growth, 2nd Edition. The MIT Press, Cambridge, Mass.

Bertram, C., Luderer, G., Pietzcker, R. C., Schmid, E., Kriegler, E., Edenhofer, O., 2015. Complementing carbon prices with technology policies to keep climate targets within reach. Nature Climate Change 5 (3), 235–239.

Borenstein, S., Davis, L. W., 2015. The Distributional Effects of U.S. Clean Energy Tax Credits. Working Paper 21437, National Bureau of Economic Research.

Bovenberg, A. L., Goulder, L. H., 1996. Optimal environmental taxation in the presence of other taxes: General- equilibrium analyses. The American Economic Review 86 (4), 985–1000.

Caldecott, B., Harnett, E., Cojoianu, T., Kok, I., Pfeiffer, A., 2016. Stranded Assets: A Climate Risk Challenge. Inter-American Development Bank.

Carney, M., 2016. Resolving the climate paradox. Arthur Burns Memorial Lecture, Bank of England, Berlin.

Chakravorty, U., Magné, B., Moreaux, M., 2006. A Hotelling model with a ceiling on the stock of pollution. Journal of Economic Dynamics and Control 30 (12), 2875–2904.

Coady, D. P., Flamini, V., Sears, L., 2015. The Unequal Benefits of Fuel Subsidies Revisited: Evidence for Developing Countries. IMF Working Paper No. 15/250.

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

Davis, S. J., Socolow, R. H., 2014. Commitment accounting of CO2 emissions. Environmental Research Letters 9 (8), 084018.

Dresner, S., Dunne, L., Clinch, P., Beuermann, C., 2006. Social and political responses to ecological tax reform in europe: an introduction to the special issue. Energy Policy 34 (8), 895–904.

Fay, M., Hallegatte, S., Vogt-Schilb, A., Rozenberg, J., Narloch, U. (Eds.), 2015. Decarbonizing development: three steps to a zero-carbon future. Climate change and development series. World Bank Group, Washington, DC.

Fischer, C., Newell, R. G., 2008. Environmental and technology policies for climate mitigation. Journal of Environmental Economics and Management 55 (2), 142–162.

Fullerton, D., Heutel, G., 2010. The general equilibrium incidence of environmental mandates. American Economic Journal: Economic Policy 2 (3), 64–89.

Fullerton, D., Heutel, G., Metcalf, G. E., 2012. Does the indexing of government transfers make carbon pricing progressive? American Journal of Agricultural Economics 94 (2), 347–353.

GAMS, 2018. GAMS - Cutting Edge Modeling. Https://www.gams.com/.

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.

Giraudet, L.-G., Quirion, P., 2008. Efficiency and distributional impacts of tradable white certificates compared to taxes, subsidies and regulations. Revue d'économie politique 118 (6), 885–914.

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.

Goulder, L. H., Schein, A. R., 2013. Carbon taxes versus cap and trade: a critical review. Climate Change Economics 04 (03), 1350010.

Grimaud, A., Lafforgue, G., 2008. Climate change mitigation policies : Are R&D subsidies preferable to a carbon tax ? Revue d'économie politique Vol. 118 (6), 915–940.

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

Guivarch, C., Hood, C., 2010. Early retirement of coal-fired generation in the transition to low-carbon electricity systems. Data and analyses.

Harrison, K., Peet, C., 2012. Historical legacies and policy reform: Diverse regional reactions to british columbia's carbon tax. BC Studies: The British Columbian Quarterly 0 (173), 97–122.

Holland, S. P., Hughes, J. E., Knittel, C. R., 2009. Greenhouse gas reductions under low carbon fuel standards? American Economic Journal: Economic Policy 1 (1), 106–146.

IEA, 2016. Policies and measures database. Tech. rep.

IPCC, 2014. Summary for policymakers. In: Climate Change 2014, Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, cambridge university press Edition. Cambridge, United Kingdom and New York, NY, USA.

Iyer, G. C., Edmonds, J. A., Fawcett, A. A., Hultman, N. E., Alsalam, J., Asrar, G. R., Calvin, K. V., Clarke, L. E., Creason, J., Jeong, M., Kyle, P., McFarland, J., Anupriya Mundra, Patel, P., Shi, W., McJeon, H. C., 2015. The contribution of Paris to limit global warming to 2°C. Environmental Research Letters 10 (12), 125002.

Jacobsen, M. R., 2013. Evaluating US Fuel Economy Standards in a Model with Producer and Household Heterogeneity. American Economic Journal: Economic Policy 5 (2), 148–187.

Jacobsen, M. R., van Benthem, A. A., 2015. Vehicle Scrappage and Gasoline Policy. American Economic Review 105 (3), 1312–38.

Jenkins, J. D., 2014. Political economy constraints on carbon pricing policies: What are the implications for economic efficiency, environmental efficacy, and climate policy design? Energy Policy 69, 467–477.

Johnson, N., Krey, V., McCollum, D. L., Rao, S., Riahi, K., Rogelj, J., 2015. Stranded on a low-carbon planet: Implications of climate policy for the phase-out of coal-based power plants. Technological Forecasting and Social Change 90, Part A, 89–102.

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

Kalkuhl, M., Edenhofer, O., Lessmann, K., 2012. Learning or lock-in: Optimal technology policies to support mitigation. Resource and Energy Economics 34, 1–23.

Kallbekken, S., Sælen, H., 2011. Public acceptance for environmental taxes: Self-interest, environmental and distributional concerns. Energy Policy 39 (5), 2966–2973.

Karp, L., Rezai, A., 2012. The Political economy of environmental policy with overlapping generations. International Economic Review 55 (3), 711-733.

Kverndokk, S., Rosendahl, K. E., 2007. Climate policies and learning by doing: Impacts and timing of technology subsidies. Resource and Energy Economics 29 (1), 58–82.

Manne, A. S., Richels, R. G., 1992. Buying greenhouse insurance: the economic costs of carbon dioxide emission limits. MIT Press, Cambridge, Mass.

McGlade, C., Ekins, P., 2015. The geographical distribution of fossil fuels unused when limiting global warming to 2c. Nature 517 (7533), 187–190.

Metcalf, G. E., 2014. Using the tax system to address competition issues with a carbon tax. National Tax Journal 67 (4), 779.

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., Bento, A. M., 2000. Tax deductions, environmental policy, and the "double dividend" hypothesis. Journal of Environmental Economics and Management 39 (1), 67–96.

Pearce, D., 1991. The role of carbon taxes in adjusting to global warming. The Economic Journal 101 (407), 938–948.

Pfeiffer, A., Hepburn, C., Vogt-Schilb, A., Caldecott, B., 2018. Committed emissions from existing and planned power plants and asset stranding required to meet the Paris Agreement. Environmental Research Letters 13 (5), 054019.

Pfeiffer, A., Millar, R., Hepburn, C., Beinhocker, E., 2016. The 2C capital stock for electricity generation: Committed cumulative carbon emissions from the electricity generation sector and the transition to a green economy. Applied Energy 179, 1395–1408.

Pigou, A., 1932. The economics of welfare. Macmillan and Co., London, 4th edition.

Rausch, S., Metcalf, G. E., Reilly, J. M., Paltsev, S., 2010. Distributional implications of alternative u.s. greenhouse gas control measures. The B.E. Journal of Economic Analysis & Policy 10 (2).

Rausch, S., Reilly, J., 2015. Carbon Taxes, Deficits, and Energy Policy Interactions. National Tax Journal 68 (1), 157–178.

Rezai, A., Foley, D., Taylor, L., 2012. Global warming and economic externalities. Economic Theory 49 (2), 329–351.

Rezai, A., Van der Ploeg, F., 2016. Intergenerational Inequality Aversion, Growth, and the Role of Damages: Occam's Rule for the Global Carbon Tax. Journal of the Association of Environmental and Resource Economists 3 (2), 493–522.

Rogelj, J., McCollum, D. L., O'Neill, B. C., Riahi, K., 2013. 2020 emissions levels required to limit warming to below 2c. Nature Climate Change 3 (4), 405–412.

Rozenberg, J., Davis, S. J., Narloch, U., Hallegatte, S., 2015. Climate constraints on the carbon intensity of economic growth. Environmental Research Letters 10 (9), 095006.

Sachs, J., Tubiana, L., Guerin, E., Waisman, H., Mas, C., Colombier, M., Schmidt-Traub, G., 2014. Pathways to Deep Decarbonization. 2014 Report, Deep Decarbonization Pathways Project.

Shearer, C., Fofrich, R., Davis, S. J., 2017. Future CO2 emissions and electricity generation from proposed coal-fired power plants in India. Earth's Future 5 (4), 2017EF000542.

Siegmeier, J., Mattauch, L., Edenhofer, O., 2018. Capital beats coal: How collecting the climate rent increases aggregate investment. Journal of Environmental Economics and Management 88, 366–378.

Sijm, J., Neuhoff, K., Chen, Y., 2006. CO2 cost pass-through and windfall profits in the power sector. Climate Policy 6 (1), 49–72.

Trebilcock, M. J., 2014. Dealing with Losers: The Political Economy of Policy Transitions. Oxford University Press.

Vogt-Schilb, A., Hallegatte, S., 2017. Climate Policies and Nationally Determined Contributions: Reconciling the Needed Ambition with the Political Economy. WIREs Energy Environ e256, dOI: 10.18235/0000714.

Vogt-Schilb, A., Meunier, G., Hallegatte, S., 2018. When starting with the most expensive option makes sense: Optimal timing, cost and sectoral allocation of abatement investment. Journal of Environmental Economics and Management 88, 210–233.

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

Weitzman, M. L., 2012. GHG Targets as Insurance Against Catastrophic Climate Damages. Journal of Public Economic Theory 14 (2), 221–244.

Williams, R. C., 2012. Setting the Initial Time-Profile of Climate Policy: The Economics of Environmental Policy Phase-Ins. In: The Design and Implementation of US Climate Policy. National Bureau of Economic Research Conference Report.

World Bank, 2016. State and Trends of Carbon Pricing 2016. The World Bank, Washington DC, USA.

Yang, Z., Bandivadekar, A., 2017. Light-duty vehicle greenhouse gas and fuel economy standards: 2017 Global update. Tech. rep., International Council on Clean Transportation, Washington.

Appendix A. Laissez-faire equilibrium (Section 3.2)

The present value Lagrangian associated to maximization of the household's utility under the capacity constraint and the capital motion law (12) is:

$$\begin{aligned} L_t = e^{-\rho t} \cdot \{ u(c_t) + \lambda_t [R_{c,t} \cdot k_{c,t} + R_{p,t} \cdot q_{p,t} - c_t - i_{p,t} - i_{c,t}] + \nu_t [i_{p,t} - \delta k_{p,t}] \\ + \chi_t [i_{c,t} - \delta k_{c,t}] + \beta_t [k_{p,t} - q_{p,t}] \} \end{aligned}$$

First order conditions read:

$$\begin{aligned} \frac{\partial L_t}{\partial c_t} = 0 \Rightarrow & \qquad u'(c_t) = \lambda_t && \text{(A.1)} \\ \frac{\partial L_t}{\partial i_{p,t}} = 0 \Rightarrow & \qquad \lambda_t = \nu_t \\ \frac{\partial L_t}{\partial i_{c,t}} = 0 \Rightarrow & \qquad \lambda_t = \chi_t \\ \frac{\partial L_t}{\partial k_{p,t}} = -\frac{d(e^{-\rho t}\nu_t)}{dt} \Rightarrow & \quad -\nu_t \delta + \beta_t = -\dot{\nu}_t + \rho \nu_t \\ \frac{\partial L_t}{\partial k_{c,t}} = -\frac{d(e^{-\rho t}\chi_t)}{dt} \Rightarrow & \lambda_t R_{c,t} - \chi_t \delta = -\dot{\chi}_t + \rho \chi_t \\ \frac{\partial L_t}{\partial q_{p,t}} = 0 \Rightarrow & \qquad \lambda_t R_{p,t} = \beta_t \end{aligned}$$

If we differentiate (A.1) with respect to time and substitute λ_t and $\dot{\lambda}_t$, we get the familiar Ramsey formula that links consumption decisions to the interest rate:

$$\frac{c_t \cdot u''(c_t)}{u'(c_t)} \cdot \frac{\dot{c}_t}{c_t} = (\rho + \delta - R_{c,t}) \tag{A.2}$$

At each time step, firms rent out all available capacities at their marginal productivity, the household observes the rental rate $R_{c,t}$ and its current consumption c_t, and chose the next step consumption level (through $\dot{c}_t$) using the Ramsey formula. This leaves one degree of freedom to the system, the choice of c_0, which is resolved by the transversality condition:

$$lim_{t \to \infty} e^{-\rho t} u'(c_t)(k_{p,t} + k_{c,t}) = 0 \tag{A.3}$$

In the laissez-equilibrium, there is no practical difference between $k_{p,t}$ and $k_{c,t}$, and this condition interprets similarly to the transversality condition in the textbook Ramsey model. It can be shown, following for instance Barro and Sala-i-Martin (2003, p. 99), that if the utility function writes $u(c) = \frac{c^{(1-\theta)}-1}{1-\theta}$, and the technological progress follows an exponential trend $A_t = e^{xt}$; then the transversality condition is equivalent to

$$\rho > (1-\theta)x \tag{A.4}$$

This condition ensures that the problem of the household (1) has a solution. Barro and Sala-i-Martin (2003) show that during the final stage, the transversality condition is equivalent to $R_{c,t} - \delta > x$. It ensures that the real returns on capital are greater than exogenous technological progress.

Appendix B. Social optimum (section 4.1)

Appendix B.1. Efficiency conditions

The present-value Lagrangian associated to the constrained maximization of social welfare (24) is:

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

All co-state variables and Lagrange multipliers are positive.

The full set of necessary first order conditions reads:

$$\frac{\partial L_t}{\partial c_t} = 0 \Rightarrow \qquad u'(c_t) = \lambda_t \tag{B.2}$$

$$\frac{\partial L_t}{\partial i_{p,t}} = 0 \Rightarrow \qquad \lambda_t = \nu_t + \psi_t \tag{B.3}$$

$$\frac{\partial L_t}{\partial i_{c,t}} = 0 \Rightarrow \qquad \lambda_t = \chi_t \tag{B.4}$$

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

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

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

$$\frac{\partial L_t}{\partial m_t} = \frac{d(e^{-\rho t}\mu_t)}{dt} \Rightarrow \qquad -\phi_t + \varepsilon \mu_t = \dot{\mu}_t - \rho \mu_t \tag{B.5}$$

Appendix B.2. Social cost of carbon

Eq. B.5 gives the evolution of μ_t. Using $\dot{\mu}_t = (\dot{\lambda}_t \tau_t + \lambda_t \dot{\tau}_t)$ from the definition of τ_t (29), (B.2), (A.2) and (36) yields:

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

We assume that GHG concentration reaches the ceiling at a date denoted t_{ss}:

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

During the final stage, $\dot{m}_t = 0 \implies G\, q_{p,t} = \varepsilon\, \bar{m}$ (11).

Before t_{ss}, $\phi_t = 0$ (32). The carbon price grows at the endogenous interest rate plus the dissipation rate of GHG before the ceiling is reached:

$$\dot{\tau}_t = \tau_t[\varepsilon + r_t] \tag{B.6}$$

Equation B.6 gives τ_t off by a multiplicative constant τ_0, which the social planer choses at the lowest value that ensures compliance with the GHG ceiling.

Given that in the long term $q_{p,t} = \varepsilon\,\bar{m}$, the transversality condition (A.3) is equivalent to $lim_{t\to\infty} e^{-\rho t} u'(c_t)(k_{c,t}) = 0$. That is, it is exactly the same transversality condition as in a textbook Ramsey model (with one type of non-polluting capital). As in the laissez-faire equilibrium, the transversality condition can be shown to require that the pure time preference of the household is large relative to the product of the absolute elasticity of the utility function and the exogenous technical progress parameter as described in (A.4).

Appendix B.3. Proof of lemma 4

THE IRREVERSIBILITY CONSTRAINT IS BINDING IN THE SHORT RUN. 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, receive the same rental rates, and installed capital is fully used (lemma 1):

$$\lim_{t\to t_0^-} q_{p,t} = k_{p,t_0} \tag{B.7}$$

$$\lim_{t\to t_0^-} \frac{\partial F}{\partial q_p}(k_{p,t}, k_{c,t}) = \frac{\partial F}{\partial k_c}(k_{p,t_0}, k_{c,t_0}) \tag{B.8}$$

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. $\psi_t = 0$. In that case, the value of clean and polluting capital are equal (B.3, B.4 $\implies \nu_t = \chi_t$), implying that the rental rate of polluting capacity is positive, $\beta_t > 0$, which in turns means that all polluting capital is used (31): $q_{p,t} = k_{p,t}$. In that case, Lemma 2 would lead to:

$$\lim_{t\to t_0^+} \frac{\partial F}{\partial q_p}(k_{p,t}, k_{c,t}) = \frac{\partial F}{\partial k_c}(k_{p,t}, k_{c,t}) + \tau_{t_0}\cdot G \tag{B.9}$$

If the GHG ceiling is binding then $\tau_{t_0} > 0$ ((B.6)). But since $\frac{\partial F}{\partial q_p}$ is a continuous function of $k_{p,t}$ and $k_{c,t}$, and $k_{p,t}$ is continuous over time, $\frac{\partial F}{\partial q_p}(k_{p,t}, k_{c,t})$ is continuous over time, thus

$$\lim_{t\to t_0^+} \frac{\partial F}{\partial q_p}(k_{p,t}, k_{c,t}) = \lim_{t\to t_0^-} \frac{\partial F}{\partial q_p}(k_{p,t}, k_{c,t}) \tag{B.10}$$

Equations B.9, B.6, and B.10 lead to a contradiction. Therefore the initial assumption is false: initially the irreversibility constraint has to be binding. □

THE IRREVERSIBILITY CONSTRAINT IS NOT BINDING IN THE LONG RUN. During the final stage, polluting capital is maintained at the maximum level compatible with stabilized GHG concentration, $k_{p,t} = \bar{m}\,\varepsilon/G$, which implies that $i_{p,t} = \delta\bar{m}\,\varepsilon/G > 0$, from which it follows that $\psi_t = 0$ and thus $\ell_t = 0$ (40).

Appendix C. Maximization of social welfare with full utilization constraint

Here, we solve a new social planner program, identical to the first best optimum, but with the additional political constraint that polluting capital should not be underused (53).

The present value Lagrangian reads:

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

First order conditions read:

$$\frac{\partial L_t}{\partial c_t} = 0 \Rightarrow \qquad u'(c_t) = \lambda_t \tag{C.1}$$

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

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

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

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

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

$$\frac{\partial L_t}{\partial m_t} = \frac{d(e^{-\rho t}\mu_t)}{dt} \Rightarrow \qquad -\phi_t + \varepsilon\mu_t = \dot{\mu}_t - \rho\mu_t$$

The complementary slackness conditions are:

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

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

$$\forall t,\ \alpha_t \geq 0 \text{ and } \alpha_t \cdot (k_{p,t} - q_{p,t}) = 0 \tag{C.4}$$

$$\forall t,\ \phi_t \geq 0 \text{ and } \phi_t \cdot (\bar{m} - m_t) = 0 \tag{C.5}$$

As before, C.5 implies that the carbon price grows at the relevant rate when the carbon budget is not saturated.

Note that in the long run, the full utilization constraint is not binding, the system tends to the same final stage as in the other sections, the transversality condition writes and interprets similarly, and imposes the same constraint on the parameters of the inter-temporal utility and exogenous technical progress (A.4).

Appendix C.1. Decentralization of the Second-Best Equilibrium with Investment Mandates or Feebates

With mandates, the household problem becomes (65).

First-order conditions for the household can be reduced to the following equations:

$$u'(c_t) = \lambda_t \tag{C.6}$$

$$\nu_t = \lambda_t + \sigma_{p,t} - \psi_t \tag{C.7}$$

$$\chi_t = \lambda_t - \sigma_{c,t} \tag{C.8}$$

$$\lambda_t R_{c,t} = (\delta + \rho)\chi_t - \dot{\chi}_t \tag{C.9}$$

$$\lambda_t R_{p,t} = \beta_t = (\delta + \rho)\nu_t - \dot{\nu}_t \tag{C.10}$$

These equations show that the mandates $s_{p,t}$ and $s_{c,t}$ impose a shadow cost and shadow subsidy on investment in polluting and clean capital respectively. Below, we show that a feebate programs that mimics those shadow values can also decentralize the constrained optimum.

Feebate. Unsurprisingly, the same transition can be obtained using a so-called feebate program that subsidizes, that is offers a rebate $\theta_{c,t}$ on investment in clean capacity and taxes, that is imposes a fee $\theta_{p,t}$ on investment in polluting capacity. With feebates, the household problem becomes 67. First order conditions for the household become:

$$u'(c_t) = \lambda_t \tag{C.11}$$

$$\nu_t = \lambda_t(1 + \theta_{p,t}) - \psi_t \tag{C.12}$$

$$\chi_t = \lambda_t(1 - \theta_{c,t}) \tag{C.13}$$

$$\lambda_t R_{c,t} = (\rho + \delta)\chi_t - \dot{\chi}_t \tag{C.14}$$

$$\lambda_t R_{p,t} = \beta_t = (\rho + \delta)\nu_t - \dot{\nu}_t \tag{C.15}$$

To decentralize the constrained social optimum, the government simply needs to set the feebate $(\theta_{p,t}, \theta_{c,t})$ such that the values of clean and polluting investment are the same as in the previous cases; that is choosing $\theta_{p,t}$ such that (C.12) is equivalent to (C.7) and choosing $\theta_{c,t}$ such that (C.13) is equivalent to (C.8). In that case, the set of equations that describe the response of the household and producer to the feebate scheme is the same as the set of equation describing their response to the mandates.

Appendix C.2. CAFE-like standards

A prominient performance standard used in actuality is the Corporate Average Fuel Economy (CAFE) standard used in the US (Yang and Bandivadekar, 2017). CAFE standards set a minimum energy efficiency on average sales of new cars in the US, which can be interpreted in our framework has a maximum average carbon intensity on new capital:

$$\frac{G\ i_{p,t}}{(i_{p,t} + i_{c,t})} \leq S_t \tag{C.16}$$

With CAFE-like standard, the household problem becomes:

$$\max_{c,i,q} \int_0^\infty e^{-\rho t} \cdot u(c_t)\, dt \tag{C.17}$$

$$\begin{aligned}
\text{subject to } & R_{c,t} \cdot k_{c,t} + R_{p,t} \cdot q_{p,t} = c_t + i_{p,t} + i_{c,t} && (\lambda_t)\\
& \dot{k}_{p,t} = i_{p,t} - \delta k_{p,t} && (\nu_t)\\
& \dot{k}_{c,t} = i_{c,t} - \delta k_{c,t} && (\chi_t)\\
& i_{p,t} \geq 0 && (\psi_t)\\
& q_{p,t} \leq k_{p,t} && (\beta_t)\\
& G i_{p,t} \leq S_t(i_{p,t} + i_{c,t}) && (\Sigma_t)
\end{aligned}$$

Firm's FOC remain unchanged, and there are not stranded assets.

The present value Lagrangian for the household reads:

$$\begin{aligned}
L_t = e^{-\rho t} \cdot \{ & u(c_t) + \lambda_t[R_{c,t} \cdot k_{c,t} + R_{p,t} \cdot q_{p,t} - c_t - i_{p,t} - i_{c,t}] + \nu_t[i_{p,t} - \delta k_{p,t}]\\
& + \chi_t[i_{c,t} - \delta k_{c,t}] + \Sigma_t[G i_{p,t} - S_t(i_{p,t} + i_{c,t})]\\
& + \psi_t \cdot i_{p,t} + \beta_t[k_{p,t} - q_{p,t}]\}
\end{aligned}$$

First order conditions for the household become:

$$u'(c_t) = \lambda_t \tag{C.18}$$

$$\nu_t = \lambda_t - \psi_t - \Sigma_t(G - S_t) \tag{C.19}$$

$$\chi_t = \lambda_t + \Sigma_t S_t \tag{C.20}$$

$$\lambda_t R_{c,t} = (\delta + \rho)\chi_t - \dot{\chi}_t \tag{C.21}$$

$$\lambda_t R_{p,t} = \beta_t = (\delta + \rho)\nu_t - \dot{\nu}_t \tag{C.22}$$

We see that standards increase the value of clean capital, while at the same time they decrease the value of polluting capital. With CAFE standards, the government has only one lever (the standard itself) to control two outputs: clean investment, and polluting investment. In general, that cannot be achieved, and the government would need an additional instrument (for instance a tax on investment or consumption) to decentralize the second-best transition using CAFE-like standards.

Appendix D. 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:

$$k_{p,t} = k_{p,0}\, e^{-\delta t}$$

Therefore, the stock of pollution follows this dynamic:

$$\dot{m} = G\, k_{p,0}\, e^{-\delta t} - \varepsilon\, m \tag{D.1}$$

The solution to this differential equation is:

$$m_t = \left(m_0 + \frac{G\, k_{p,0}}{\delta - \varepsilon} \right)\, e^{-\varepsilon t} - \frac{G\, k_{p,0}}{\delta - \varepsilon}\, e^{-\delta t}$$

This function reaches its maximum m_{max} at the date t_{max} when $\dot{m} = 0$. The maximum date is thus

$$t_{max} = -\frac{1}{\delta}\, \ln(\frac{\varepsilon\, m_{max}}{G\, k_{p,0}})$$

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

$$\bar{m} = \left(m_0 + \frac{G\, k_{p,0}}{\delta - \varepsilon} \right)\, e^{\frac{\varepsilon}{\delta} \ln(\frac{\bar{m}\, \varepsilon}{G\, k_{p,0}})} - \frac{G\, k_{p,0}}{\delta - \varepsilon}\, e^{\ln(\frac{\bar{m}\, \varepsilon}{G\, k_{p,0}})}$$

This can be rewritten:

$$\bar{m} = \left[\left(m_0 + \frac{G\, k_{p,0}}{\delta - \varepsilon} \right) \left(\frac{\varepsilon}{G\, k_{p,0}} \right)^{\frac{\varepsilon}{\delta}} \left(\frac{\delta - \varepsilon}{\delta} \right) \right]^{\frac{\delta}{\delta - \epsilon}}$$

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

Since realistic values for the natural decay of atmospheric GHG ε, less than 0.4% per year (e.g. Rezai et al., 2012), are negligible with respect to capital depreciation $\varepsilon \ll \delta$, the previous relation can be approximated by $\bar{m} = m_0 + \frac{G\, k_{p,0}}{\delta}$, which is also simply the maximum of the solution of (D.1) when $\varepsilon = 0$.

Appendix E. Numerical model

Appendix E.1. Approach

To illustrate this paper, we use simulations of the social planner problems 24 and 53 solved in GAMS. The General Algebraic Modeling System (GAMS) is a high-level language designed to solve optimization problems directly from a description of the social planner problem (GAMS, 2018). Since the purpose is to illustrate the article, we chose a calibration of the model that yields interesting figures (see below). To obtain Figure 1 and Figure 6, we varied systematically *betta* and *mbar* (which correspond to G and $\bar{m}$ in the manuscript) in the code below.

Appendix E.2. Code

```
Parameters
dt                "length of time period"          /1/
;

Sets
t                 "Time periods"       /1*200/
```

```
tfirst(t)        "first period"
;

tfirst(t) = yes$(ord(t) eq 1);

Parameters
rho        "discount rate" /0.01/
betta    "co2 intensity of capital" /0.005       /
kb0        "initial stock of brown capital" /10  /
delta      "capital depreciation" /0.05    /
epsilon    "carbon depreciation"  /0.003   /
g              "exogenous growth rate" /0.02     /
mbar     "carbon budget (Tt)" /1.3   /
m0           "Init cumulative emissions (Tt)" /0.5     /
alpha    "brown k productivity" / 0.303    /
gama     "green k productivity" / 0.0303   /
A0           "Init A" / 40     /
;


Positive Variables
c(t)              "consumption"
ig(t)             "green investment"
qb(t)                   "brown used capital"
;

variable
ib(t)             "brown investment"
e(t)                    "emissions"
kg(t)                   "green capital"
kb(t)                   "brown capital"
m(t)                    "carbon stock"
y(t)                    "output"
A(t)                    "total productivity"
utot              "utility"
;

Equations
eq_set_kb0              "cond init"
eq_set_qb0              "cond init"
eq_set_kg0              "cond init"
eq_set_m0               "cond init"
eq_set_A0               "cond init"

eq_set_mbar              "carbon budget constraint"
eq_set_underusedcapital  "under utilization"
eq_set_fullusedcapital   "full utilization"
eq_set_irreversibility   "irreversibility"

eq_cal_totutil        "tot discounted utility"
eq_cal_output         "net output"
eq_cal_emissions            "emissions from brown capital"
eq_cal_conso                "budget equation"

eq_dyn_kb                   "dyn brown capital"
eq_dyn_kg                   "dyn green capital"
eq_dyn_m                    "dyn carbon stock"
```

```
eq_dyn_A                    "growth"
;

eq_set_A0(tfirst)..
      A(tfirst) =e= A0;

eq_set_m0(tfirst)..
      m(tfirst) =e= m0;

eq_set_kb0(tfirst)..
      kb(tfirst) =e= kb0;

eq_set_qb0(tfirst)..
      qb(tfirst) =e= kb(tfirst);

eq_set_kg0(tfirst)..
      kg(tfirst) =e=    kb(tfirst)*(gama/alpha);

eq_set_mbar(t)..
      m(t) =l= mbar;

eq_set_underusedcapital(t)..
      qb(t) =l= kb(t);

eq_set_fullusedcapital(t)..
      qb(t) =e= kb(t);

eq_set_irreversibility(t)..
      ib(t) =g= 0;

eq_dyn_m(t-1)..
      m(t) =e= m(t-1) + e(t-1)*dt  - epsilon*m(t-1)*dt;

eq_cal_output(t)..
            y(t) =e= A(t)*((0.00001+qb(t))**alpha)*(kg(t)**gama);

eq_cal_emissions(t)..
      e(t) =e= betta*qb(t);

eq_cal_conso(t)..
      c(t) =l= y(t) - ib(t) - ig(t);

eq_cal_totutil..
      utot =e= sum((t), log(1+c(t))/  ( exp(rho*ord(t)*dt))   )*dt;

eq_dyn_kb(t-1)..
      kb(t) =e= kb(t-1) + ib(t-1)*dt - delta*kb(t-1)*dt;

eq_dyn_kg(t-1)..
      kg(t) =e= kg(t-1) + ig(t-1)*dt - delta*kg(t-1)*dt;

eq_dyn_A(t-1)..
      A(t) =e= A(t-1) + g*A(t-1)*dt;

model baseline/ eq_set_kb0 ,eq_set_qb0 ,eq_set_kg0 ,eq_set_m0 ,
    eq_set_A0 ,eq_set_fullusedcapital ,eq_cal_totutil ,eq_cal_output ,
    eq_cal_emissions ,eq_cal_conso ,eq_dyn_kb ,eq_dyn_kg ,eq_dyn_m ,
```

```
    eq_dyn_A ,eq_set_irreversibility /;

model under / eq_set_kb0,eq_set_qb0,eq_set_kg0,eq_set_m0,eq_set_A0,
    eq_set_mbar,eq_set_underusedcapital,eq_cal_totutil,eq_cal_output,
    eq_cal_emissions,eq_cal_conso,eq_dyn_kb,eq_dyn_kg,eq_dyn_m,eq_dyn_A
    ,eq_set_irreversibility/;

model full / eq_set_kb0,eq_set_qb0,eq_set_kg0,eq_set_m0,eq_set_A0,
    eq_set_mbar,eq_set_fullusedcapital,eq_cal_totutil,eq_cal_output,
    eq_cal_emissions,eq_cal_conso,eq_dyn_kb,eq_dyn_kg,eq_dyn_m,eq_dyn_A
    ,eq_set_irreversibility/;


solve full using nlp maximizing utot

solve under using nlp maximizing utot
```