Environmental policy in the presence of bias and uncertainty*

Maria Arvaniti† Niko Jaakkola‡

April 2025

Abstract

We develop a theoretical model of the European Union emissions trading scheme (EU ETS) in which participants design collectively optimal policy, understanding that they will face heterogeneous political pressure to conduct more or less ambitious policy *ex post*. Member states can commit to an *ex ante* minimum requirement on complementary emission reduction technologies. If policy on these complementary technologies is designed at the Union level, this minimum requirement will also take into account the redistributive effect through permit allocation and the global externality. However, asymmetric information still distorts the emission cap away from the constrained-efficient outcome.

Keywords: climate change; bias; uncertainty; commitment; flexibility

JEL Classification: Q54; D70; D50

*We thank ....

†maria.arvaniti@unibo.it, University of Bologna, Department of Economics, Italy, Affiliate Member CESifo

‡niko.jaakkola@unibo.it, University of Bologna, Department of Economics, Italy.

1 Introduction

The European Union emissions trading scheme (EU ETS) remains a rare success in the global fight against climate change. The carbon pricing scheme covers nearly half of European Union emissions, or about 5% of global emissions, and the price imposed on emitting a tonne of carbon has been substantial, especially in recent years. Even when carbon prices were lower, in the period 2005-12, the system achieved a meaningful reduction in overall emissions (Colmer et al., 2024).

One possible reason for the success of EU ETS is that it partially delegates the management of the policy, including the setting of emission permit trajectories to a supranational and technocratic institution—the European Commission. International and supranational institutions are thought to be able to act in a more long-termist manner, as they are insulated from the day-to-day pressures of electoral politics leading to more short-termist policy decisions (Leruth, 2024). At the same time, carbon emissions are better regulated at a European level rather than at member state level, as collective governance is able to take into account the Union-wide benefits of reducing climate pollutants.

Two other key features characterise climate policy. First, changing local economic or political circumstances mean that the *ex post* effect of policies is liable to diverge from that envisioned *ex ante*. One salient example is how the cutting off of supplies of natural gas to Europe associated with Russia's war in Ukraine led to Germany—traditionally a frontrunner in European climate policy—reopening highly polluting coal-fired power plants. Second, the EU ETS is not the sole climate policy which the member states rely on. Again, Germany provides the example, in the shape of the high feed-in tariffs employed to promote the construction of renewable power generation (with great success).

In this paper, we adapt a model of fiscal rules in a supranational setting (Halac and Yared, 2018) to study how collective and decentralised forms of climate policymaking interact, when the policy has to take into account both the short-termism of policymakers and uncertainty regarding the policies' effects. The central tension in our model is between commitment and flexibility (Amador et al., 2006). In particular, the policymakers would like to collectively commit to policies *ex ante*, as they know they will be tempted to relax policies *ex post* due to political pressures. At the same time, full commitment is undesirable as circumstances may unexpectedly increase or decrease the value of carbon emissions to the national decisionmaker.

In our model, much as in the real European Union, the policymakers jointly commit to an emissions trading scheme. This scheme will be costly to high emitting countries, as they end up being net payers, transferring resources to less carbon-intensive countries. The countries do not know *ex ante* their benefit of carbon emissions, reflecting e.g.

the local environmental co-benefits of reducing carbon emissions, or perhaps more realistically the willingness-to-pay to reduce emissions which the national political process produces. However, governments has access to an alternative emission reduction technology, at a convex cost. This could reflect complementary policies to reduce emissions. However, it could also be interpreted as a costly negotiation to obtain a higher share of the emission permit allocation, so that the cost would reflect trade-offs made regarding the government's other political objectives vis-a-vis their EU partners.[1]

We show that policymakers trade off commitment and discretion, as they are aware of the pressure they are later under to invest less in pollution reductions. The particular means of commitment is a minimum complementary investment rule, which requires those governments under the most political pressure to invest at least a given amount into complementary emission reductions. If this minimum investment rule is designed at the Union level, in a coordinated manner, then the minimum investment requirement takes into account the effect of the rule on the carbon permit price, which has three kinds of effects on member states. First, a higher permit price implies more redistribution of permit revenues from high-emitting countries to low-emitting countries. Second, a higher permit price improves the incentives of member states to invest in emission reduction technologies; and third, also motivates domestic firms to reduce emissions directly. Given that the aggregate emission cap is fixed, these effects change the *distribution* of emissions between member states, and the optimal rule is set to account for these features. The first two effects are effectively the "redistributive" and "disciplining" effects of Halac and Yared (2018), but the third is novel. However, in our paper, there is also an "externality" effect relating to the distribution of emissions and thus of the damage (whether perceived or physical) in the member states, which is also taken into account.

We also consider the case in which the emission cap itself is determined optimally, in a coordinated manner, at the same time as the coordinated rule. In this situation, the information friction preventing the imposition of true type-specific complementary investments also implies that the second-best emission cap deviates from the constrained efficient cap.

As discussed above, the closest paper to our work is Halac and Yared (2018), whose model of an optimal fiscal rule (using a mechanism design approach) we build on. However, our institutional context is subtly different. First, we expressly consider the control of a pollutant imposing an externality or all Union members, controlled by an emission trading scheme. The choice of the overall quantity of emissions is thus endogenous in our model (unlike net debt, which is zero in the baseline case of Halac and Yared, 2018). Second, redistributive concerns in our model relate to the transfers resulting from the

[1]Currently, 57% of EU ETS permits are auctioned, with the rest given as free allocation. However, only 90% of the auctioned permits are allocated to member states according to historical emissions; 10% are allocated separately "for the purposes of solidarity, growth and interconnections within the Union", suggesting flexibility in actual permit allocation based on political circumstances (The European Commission, 2020).

carbon pricing scheme and also the heterogenous local benefit of complementary emission reductions. We also think our application of an emissions trading system is a natural setting for the approach of Halac and Yared (2018): the institution is well developed and clearly has the competence and authority to monitor carbon emissions and collect permits, unlike in the case of fiscal rules, which have often been observed mainly in the breach (Bilbiie et al., 2021).

We are also close to Martimort and Sand-Zantman (2016), who similarly consider a mechanism design approach to international climate agreements. However, in their model the focus is on truly international settings, in which the agreement must satisfy participation constraints for governments, which could otherwise withdraw from the agreement. We assume (realistically) that exiting the European Union over climate policy is far too costly to contemplate, so that the issue is not participation, but joint decisionkaing when the member states face a dilemma between commitment and flexibility.

There is of course a substantial literature of international environmental agreements in static settings (Hoel, 1992; Carraro and Siniscalco, 1993; Barrett, 1994). We depart from the former by not considering participation constraints and by assuming that the members of the climate coalition (the EU) face ex post heterogeneity and temptation. We do not model dynamic considerations (Harstad, 2012, 2016).

2 The Model

The economy consists of a continuum of countries/governments of unit mass who undertake activities to mitigate climate damages through emissions abatement and green investment. In each country i, total energy y_i is given by green energy (renewables) and emissions which we assume are perfect substitutes so that $y_i = g_i + e_i$. Total energy consumption implies strictly increasing and concave country-specific benefits from the productive activities of a representative firm, $B(y_i)$.

Emissions by country i generate two kinds of damages. First, there is a local damage $d(e_i)$, external to emitting firms but contained within the country. These emissions also contribute to global damage (in all countries) $D(E)$, where E are global emissions. Both function d and D are strictly increasing and convex.

We assume that countries participate in an international permit market. The number of permits issued to the representative domestic firm in country i amounts to ω_i.[2] As firms in all countries require emission permits in an amount equal to the emissions e_i they produce, global emissions are given by the sum of the emission permits issued, i.e., $E = \Omega$ which are exogenous for now. We will assume that the emission cap Ω is divided equally between the ex-ante symmetric countries so that $\omega_i = \omega$ for all i.

[2]The method of allocating permits has no bearing on our results. Grandfathering the permits and auctioning them are equivalent in our setting.

Since permits are traded internationally a firm can generate additional profits by selling permits to the firm in another country or reducing the compliance cost by buying permits from abroad. Thus, the profits of the representative firm in country i read:

$$\pi_i(e_i; g_i, p, \omega_i) = B(e_i + g_i) + p \times (\omega_i - e_i) \tag{2.1}$$

where p is the price of permits in an international market. For the later analysis, we define $T_i = p \times (\omega_i - e_i)$ to be the (sign-unconstrained) financial transfer through the permit market. Taking as given the level of green investment and the permit price, profit maximisation and equilibrium in the energy market require equalisation of marginal benefits emissions with the international permit price as

$$p = B'(e_i + g_i) \tag{2.2}$$

which gives us the optimal level of emissions $\tilde{e}^*(g, p)$. As fossil and green energy are perfect substitutes we have $\tilde{e}^*_g = -1$, and of course, fossil energy use decreases in the permit price so that $\tilde{e}^*_p = B''(\tilde{e}^*(g, p) + g)^{-1} < 0$.

Countries, taking into account the emission choice of the representative firm, choose the level of green investment. Importantly, they differ in their environmental preference, ϕ_i: this is the weight on consumption utility relative to environmental damages. The types ϕ_i are independently drawn from the cumulative distribution $F(\phi)$ with support $\Phi = [\underline{\phi}, \bar{\phi}]$, $\underline{\phi} > 0$, with atomless everywhere positive density $f(\phi) = F'(\phi)$ (continuously differentiable function) and $\mathbb{E}(\phi) = 1$. Importantly, this preference parameter or taste shock is private information of the government, therefore international and domestic environmental policy cannot be conditioned on countries' types. As we will show this introduces a preference for flexibility in our model.

Prior to the realisation of ϕ, and subject to the constraints imposed by the international and/or domestic environmental policy (represented by the permit price p and the permit endowment ω_i), a country/government would choose investment g_i to maximise

$$\mathbb{E}_\phi\left[\phi(\pi(\tilde{e}^*(g_i, p); g_i, p, \omega_i) - c(g_i)) - (d(\tilde{e}^*(g_i, p)) + D(E))\right] \tag{2.3}$$

, where the expectation is taken with respect to ϕ and E, are global emissions which a country takes as given. The cost of investing in green technology is $c(g_i)$ with $c' > 0$ and $c'' > 0$. Notice that a higher value of ϕ implies a higher weight on the consumption utility and a higher marginal benefit from energy consumption, relative to the environmental cost.

Instead, a country's objective after the realisation of the shock/type is

$$\phi_i(\pi(\tilde{e}^*(g_i, p); g_i, p, \omega_i) - c(g_i)) - \beta(d(\tilde{e}^*(g_i, p) + D(\mathbb{E}(E))) \tag{2.4}$$

where $\beta \in (0, 1]$ introduces a bias towards the emissions generating productive activity: a country's welfare after the realisation of its type overweights the importance of the benefits from energy consumption compared to its welfare prior to the realisation. The government's optimal policy is then obtained from the first-order condition as $g^*(\phi_i, p, \omega_i)$. We also denote $e^*(\phi, p) = \tilde{e}^*(g^*(\phi, p, \omega), p)$. In what follows, we often suppress the argument ω_i for notational clarity. Government green spending decreases in the type, i.e. $g^*_\phi < 0$; but the marginal effect of the permit price on government green spending is of ambiguous sign.[3]

This divergence in the objectives creates a preference for commitment and in a framework with uncertainty as ours it introduces a trade-off between commitment and flexibility. To see this, notice that in the absence of any constraints imposed by international and/or domestic policy, a country as a function of its type would set investment ex-ante to satisfy.

$$\phi_i B'(\tilde{e}^*(g_i, p) + g_i^*) + d'(e_i^*) = \phi_i c'(g_i^*) \tag{2.5}$$

where the left-hand side gives the benefit of green investment, i.e. a marginal increase in output and a marginal displacement of polluting fossil fuel use; and the right-hand side gives the marginal cost of green investment. However, ex-post, the biased government will instead choose investment to satisfy.

$$\phi_i B'(\tilde{e}^*(g_i, p) + g_i^*) + \beta d'(e_i^*) = \phi_i c'(g_i^*) \tag{2.6}$$

and would therefore underinvest. Hence, the ex-ante optimum in (2.5) can never be implemented with full flexibility when countries are biased and cannot be implemented with full commitment either: an investment plan cannot be conditioned on a country's type, therefore the ex-ante optimum cannot be implemented by fully committing a government to a contingent plan. These two frictions—the bias and private information of types—are at the heart of the trade-off between commitment and flexibility. Notice that in our setup, there is a third friction relating to the externality: a country/government takes into account only the damages from domestic emissions, and ignores the externality imposed on all countries through global emissions E. A global social planner/government would instead take the externality into account and would implement a lower level of emissions (a higher level of investment) ceteris paribus. The interplay between all these forces makes the design of the optimal international and domestic environmental policy a complicated task that we will explore in the next sections.

[3] We have
$$g^*_\phi(\phi, p) = -\frac{\beta d'/\phi}{\phi c'' + \beta d''}, \qquad g^*_p(\phi, p) = \frac{\phi + \beta d''/B''}{\phi c'' + \beta d''}.$$
.

Finally, equilibrium in permit markets requires

$$\int_{\underline{\phi}}^{\bar{\phi}} e(\phi, p) f(\phi) \mathrm{d}\phi = \Omega = E. \tag{2.7}$$

It is of course the permit price which will adjust to ensure this holds. The equilibrium permit price $p^*(E)$ satisfies

$$\frac{\mathrm{d}p^*}{\mathrm{d}E} = \left(\int_{\underline{\phi}}^{\bar{\phi}} e_p(\phi, p) f(\phi) \mathrm{d}\phi \right)^{-1}.$$

2.1 Constrained efficiency

As a brief aside, before analysing the decentralised problem further, we derive as a benchmark a constrained efficient outcome. We assume that the social planner can condition green investment directly on the type and can set the parameters of the emission trading system, choosing the permit quantity and a permit price. The permit price will then drive the representative firms' fossil fuel usage choices. We focus on this constrained efficient outcome, rather than full efficiency because an important element of our analysis concerns income transfers associated with emission trading, and full efficiency would eliminate this aspect from consideration.[4] The social planner maximises expected utility:[5]

$$\begin{aligned} \max_{g(\phi), E, p} \int_{\underline{\phi}}^{\bar{\phi}} & \left[\phi(B(\tilde{e}^*(g(\phi), p) + g(\phi, p)) + p(E - \tilde{e}^*(g(\phi), p)) - c(g(\phi, p))) \right. \\ & \left. -(d(\tilde{e}^*(g(\phi), p)) + D(E)) \right] f(\phi) \mathrm{d}\phi, \end{aligned}$$

subject to $\tilde{e}^*(g(\phi), p)$ being given by $B'(\tilde{e}^*(g^*) + g^*) - p = 0$; and the constraint

$$\int_{\underline{\phi}}^{\bar{\phi}} \tilde{e}(g(\phi), p) f(\phi) \mathrm{d}\phi = E.$$

Denoting the Lagrange multiplier of this constraint by μ, the necessary conditions

[4] Rather, because utility is linear in money but with marginal utility varying across countries, allowing the social planner to make unrestricted transfers leads to nonsensical outcomes. The other alternative is to ignore redistributive concerns from the social welfare function, but then the planner's objective deviates from the aggregate welfare of the decentralised problem.

[5] Note that since countries are ex-ante identical, the distribution of realised types across countries is the same as the distribution of types for each country.

for this problem are:

$$\phi B'(\tilde{e}^* + g) + d'(\tilde{e}^*) + \mu = \phi c'(g), \qquad \text{for all } \phi; \tag{2.8}$$

$$\mu = D'(E) - \mathbb{E}\left[\phi\right] p, \tag{2.9}$$

$$\mu \int_{\underline{\phi}}^{\bar{\phi}} \tilde{e}_p^* f(\phi) \mathrm{d}\phi = \int_{\underline{\phi}}^{\bar{\phi}} \left[\phi(E - \tilde{e}^*) - d'(\tilde{e}) \tilde{e}_p^*\right] f(\phi) \mathrm{d}\phi. \tag{2.10}$$

Equation (2.8) shows that the planner's optimal investment satisfies a condition otherwise identical to that of the individual government's ex-ante decentralised FOC, but taking into account an additional benefit of government investment μ. Equation (2.9) shows that this benefit is related to the wedge between the permit price and the marginal global externality cost; if the planner chooses these two quantities to be unequal, then it becomes worthwhile to adjust green investment away from the first best in response. Finally, equation (2.10) demonstrates why the permit price might not be equal to the global externality cost: (i) because of distributional concerns (the first term in the integral on the RHS), and (ii) because of local externality concerns (the second term in the same integral).

Putting the latter two conditions together, we obtain

$$\mathbb{E}\left[\phi\right] p = D'(E) + \frac{\mathrm{d}p^*}{\mathrm{d}E} \left(\int_{\underline{\phi}}^{\bar{\phi}} d'(\tilde{e}) \tilde{e}_p f(\phi) \mathrm{d}\phi - \int_{\underline{\phi}}^{\bar{\phi}} \phi(E - \tilde{e}) f(\phi) \mathrm{d}\phi \right) \tag{2.11}$$

Thus, the constrained efficient price of carbon, after being weighted with the average marginal utility of money, takes into account the global externality, the weighted local externality (where the weight reflects the effectiveness of carbon prices in cutting emissions), as expected; however, the impact of these taxes is attenuated by concerns that a high carbon price tends to transfer monetary resources away from countries who put a high weight on these resources.[6]

Thus, the constrained efficient outcome tends to feature more investment than the decentralised outcome because of (1) lack of the bias term β (compare equations (2.8) and (2.6) to see this) and (2) the taking into account of the global externality; but this is at least partly offset by (3) the distributional concerns associated with the permit market.

2.2 Uncoordinated Policy

We start by analysing the uncoordinated investment policy. Each country/government independently chooses and commits to an investment rule, ϕ^*, to maximise its expected

[6]Note that the third term is negative; $e_p^* < 0$, and the integral takes the permit market constraint, equal to zero, and puts more weight ϕ on countries with high emissions, for whom the term $E - e^*$ tends to be more negative or less positive.

welfare taking as given the total level of permits/emissions, the permit price and the emission choice of the representative firm and its ex-post optimising behaviour:

$$\int_{\underline{\phi}}^{\phi^*} [\phi(B(y^*(\phi,p)) + p(\omega - e^*(\phi,p)) - c(g^*(\phi,p))) - (d(e^*(\phi,p))) + D(E))] f(\phi)d\phi$$
$$+ \int_{\phi^*}^{\bar{\phi}} [\phi(B(y^*(\phi^*,p)) + p(\omega - e^*(\phi^*,p)) - c(g^*(\phi^*,p)) - (d(e^*(\phi^*,p)) + D(E))] f(\phi)d\phi \tag{2.12}$$

In choosing the optimal investment rule ϕ^* (and the corresponding investment floor), the government realises that all types below the chosen level ϕ^* will have full discretion and will choose their investment optimally given their type. Types above ϕ^* would prefer to choose $g < g^*(\phi^*, p)$ but the rule prevents this; their optimal action is then to choose $g^*(\phi^*, p)$.

It is easy to show that the optimal cut-off value ϕ^* is interior and is given by the following first-order condition.

$$\int_{\phi^*}^{\bar{\phi}} \left[\phi(B'(y^*(\phi^*,p)y^*_{\phi} - p(\phi^*,E)e^*_{\phi} - c'(g^*(\phi^*,p)g^*_{\phi}) - d'(e^*(\phi^*,p))e^*_{\phi}\right] f(\phi)d\phi = 0$$

which simplifies to

$$\int_{\phi^*}^{\bar{\phi}} \left[\left(\phi(p(\phi^*,E) - c'(g^*(\phi^*,p)\right) g^*_{\phi} + d'(e^*(\phi^*,p))g^*_{\phi}\right] f(\phi)d\phi = 0 \tag{2.13}$$

The above condition implies that the optimal uncoordinated investment rule is set so that the average distortion above the cutoff ϕ^*_u is 0: there exists $\hat{\phi} > \phi^*_u$ such that for $\phi \in [\phi^*_u, \hat{\phi}]$

$$\phi c'(g^*(\phi^*,p) > \phi p(\phi^*,E) + d'(e^*(\phi^*,p))$$

and hence the government overinvests (and underemmits) relative to its ex-ante optimum while for $\phi \in (\hat{\phi}, \bar{\phi}]$

$$\phi c'(g^*(\phi^*,p) < \phi p(\phi^*,E) + d'(e^*(\phi^*,p))$$

and a government underinvest relative to its ex-ante optimum. Therefore, the cutoff ϕ^*_u defines an optimal trade-off between more discretion for the types that underinvest and more flexibility for the types that overinvest.

Substituting equation (2.6), the above expression simplifies further to

$$\int_{\phi^*}^{\bar{\phi}} \left(1 - \beta \frac{\phi}{\phi^*}\right) g^*_{\phi} d'(e^*(\phi^*, p) f(\phi) d\phi = 0 \tag{2.14}$$

which after some algebraic manipulation gives us our first result.

Proposition 1 . *For any price p, the optimal uncoordinated investment rule specifies a cutoff value ϕ^* that satisfies*

$$\frac{\mathbb{E}[\phi|\phi \geq \phi^*]}{\phi^*} = \frac{1}{\beta} \tag{2.15}$$

We see that the optimal uncoordinated investment depends on the type and the level of bias β but is independent of the functional forms and the price of permits. [7] If $\beta = 1$, equation (2.15) implies that $\phi^* = \overline{\phi}$ and full flexibility: when governments are not biased, there is no reason to restrict their choices. If $\beta \leq \underline{\phi}$, then $\phi^* \leq \underline{\phi}$: in this case, the degree of bias is so high that the government prefers to forgo all flexibility and bunch all types at the same level of investment. Finally, for $\beta \in [\underline{\phi}, 1]$, the optimal uncoordinated rule balances the tradeoffs between commitment and flexibility and defines an optimal level $\phi^* \in (\underline{\phi}, \overline{\phi})$.

Additionally, we can show the following:

Corollary 1 . *If $F(\phi)$ satisfies*

$$\frac{d \log \mathbb{E}[\phi|\phi \geq \phi^*]}{d \log \phi^*} < 1 \tag{2.16}$$

for all $\phi^* \leq \overline{\phi}$, then ϕ^* is increasing in β.[8]

2.3 Coordinated policy

We now turn to the choice of coordinated policies. In this case, the countries get together ex-ante to jointly agree on a coordinated investment rule ϕ^{**} and on the quantity of emission permits E^{**}. They still maximise expected ex-ante social welfare given by equation (2.12); the key difference is that the countries consider the permit price to be endogenous.

We thus need to understand how the investment rule ϕ^{**} and total permit quantity E affect the permit price. Assuming a strictly positive permit price, we have

$$\int_{\underline{\phi}}^{\phi^{**}} e^*(\phi, p) f(\phi) \mathrm{d}\phi + \int_{\phi^{**}}^{\overline{\phi}} e^*(\phi^{**}, p) f(\phi) \mathrm{d}\phi = E,$$

which yields the equilibrium price $p^*(\phi^{**}, E)$. Implicit differentiation then shows that the permit price increases with the minimum investment threshold and decreases with

[7]This follows from the multiplicative nature of shocks.

[8]Condition (2.16) is satisfied by all log-concave densities which includes several familiar distributions, (Bagnoli and Bergstrom 2005)

the total permit allocation, i.e. $p^*_{\phi^{**}} > 0$, $p^*_E < 0$. The actual expressions for these partial derivatives are complicated and uninformation, but we note that $p^*_{\phi^{**}}/p^*_E = \int_{\phi^{**}}^{\overline{\phi}} g^*_{\phi^{**}}(\phi, p) f(\phi) \mathrm{d}\phi$.

2.3.1 Exogenous E

We first derive the optimal investment rule for any permit quantity E. Countries jointly choose a cut-off value ϕ^{**} to maximise

$$\begin{aligned}
&\int_{\underline{\phi}}^{\phi^{**}} [\phi(B(y^*(\phi, p^*(\phi^{**}, E))) + p^*(\phi^{**}, E)(\omega - e^*(\phi, p^*(\phi^{**}, E))) \\
&-c(g^*(\phi, p^*(\phi^{**}, E)))) - (d(e^*(\phi, p^*(\phi^{**}, E)))) + D(E))]\, f(\phi) d\phi \\
&+ \int_{\phi^{**}}^{\overline{\phi}} [\phi(B(y^*(\phi^{**}, p^*(\phi^{**}, E))) + p(\omega - e^*(\phi^{**}, p^*(\phi^{**}, E))) \\
&- c(g^*(\phi^{**}, p^*(\phi^{**}, E)) - (d(e^*(\phi^{**}, p^*(\phi^{**}, E))) + D(E))]\, f(\phi) d\phi
\end{aligned} \tag{2.17}$$

taking into account the ex-post behaviour of countries and the permit market clearing condition. The first-order condition yields

$$\begin{aligned}
&p^*_\phi \left[\int_{\underline{\phi}}^{\phi^{**}} \phi(\omega - e^*(\phi, p^*(\phi^{**}, E))) f(\phi) d\phi + \int_{\phi^{**}}^{\bar{\phi}} \phi(\omega - e^*(\phi^{**}, p^*(\phi^{**}, E))) f(\phi) d\phi \right] + \\
&p^*_\phi \left[\int_{\underline{\phi}}^{\phi^{**}} d'(e^*(\phi, p^*(\phi^{**}, E))((1-\beta) g^*_p - e^*_p) f(\phi) d\phi + \int_{\phi^{**}}^{\bar{\phi}} d'(e^*(\phi^{**}, p^*(\phi^{**}, E))((1-\beta \frac{\phi}{\phi^{**}}) g^*_p - e^*_p) f(\phi) d\phi \right] \\
&+ \int_{\phi^{**}}^{\bar{\phi}} \left(1 - \beta \frac{\phi}{\phi^{**}}\right) g^*_\phi d'(e^*(\phi^*, p^*(\phi^{**}, E)) f(\phi) d\phi = 0
\end{aligned} \tag{2.18}$$

The first row of equation (2.18) is related to the direct price effects going through the permit market and the implicit transfers to countries. The second row is associated with the indirect price effects going through the optimal investment and emissions choices while the third row relates to the direct effects of the cut-off rule as in the case of the uncoordinated actions.

Proposition 2 . *For any level of E, the optimal coordinated investment rule specifies a* *cut-off value ϕ^{**} and an associated $p^*(\phi^{**}, E)$ which satisfies*

$$\frac{\mathbb{E}[\phi|\phi \geq \phi^{**}]}{\phi^{**}} = \frac{1}{\beta} + \frac{p^*_\phi}{\beta d'(e^*(\phi^{**}, p^*(\phi^{**}, E)))g^*_\phi(1 - F(\phi^{**}))} [\rho(\phi^{**}) + \lambda(\phi^{**}, \beta)$$
$$- \int_{\underline{\phi}}^{\phi^{**}} d'(e^*(\phi, p^*(\phi^{**}, E)))e^*_p f(\phi)d\phi - \int_{\phi^{**}}^{\bar{\phi}} d'(e^*(\phi^{**}, p^*(\phi^{**}, E)))e^*_p f(\phi)d\phi\Bigg] \quad (2.19)$$

where

$$\rho(\phi^{**}) = \left[\int_{\underline{\phi}}^{\phi^{**}} \phi(\omega - e^*(\phi, p^*(\phi^{**}, E)))f(\phi)d\phi + \int_{\phi^{**}}^{\bar{\phi}} \phi(\omega - e^*(\phi^{**}, p^*(\phi^{**}, E)))f(\phi)d\phi\right] \leq 0 \quad (2.20)$$

and

$$\lambda(\phi^{**}, \beta) = \Bigg[\int_{\underline{\phi}}^{\phi^{**}} d'(e^*(\phi, p^*(\phi^{**}, E))(1 - \beta)g^*_p f(\phi)d\phi$$
$$+ \int_{\phi^{**}}^{\bar{\phi}} d'(e^*(\phi^{**}, p^*(\phi^{**}, E)))(1 - \beta\frac{\phi}{\phi^{**}})g^*_p f(\phi)d\phi\Bigg] \lesseqgtr 0 \quad (2.21)$$

We can see that the optimal coordinated investment rule differs from the uncoordinated by the second and third terms that appear in equation (2.19) and are associated with the direct and indirect effects of the price on the optimal choices and which are not internalised by the uncoordinated rule.

The term ρ can be thought of as a redistributive effect and it is negative. Higher permit prices imply a higher cost for the permit buyers and since countries with high ϕ are the ones who emit more (and invest less), they will benefit from a price reduction more than the low ϕ countries. Moreover, it is exactly those countries that put more weight on these resources hence the central planner who cares for aggregate (average) welfare will put a higher weight on them and will have an incentive to reduce the permit price to redistribute resources from the lower to the higher types through the permit market. To see this more clearly, assume that condition (2.16) holds so that the LHS of (2.19) is decreasing in ϕ^{**}. Then, holding everything else fixed, we see that a higher value of ρ (in absolute terms) implies a lower value of ϕ^{**}: there is a decrease in the level of flexibility for the sake of decreasing the permit price and transfer resources from the ones who buy fewer permits (or sell) to those who have to buy a lot of permits. [9]As noted in the constrained efficiency discussion, the redistributive effect is present even in the case of no present bias ($\beta = 1$) since we are in a set up of heterogeneous countries with allocations that cannot be type contingent.

Next, we consider the disciplining effect of the permit price, λ, which captures the

[9]Note that $\frac{p^*_\phi}{\beta d'(e^*(\phi^*, p^*(\phi^{**}, E))g^*_\phi(1-F(\phi^{**}))} < 0$.

fact that the permit price affects the level of green spending (investment) and it can be positive or negative depending on the sign of g_p^*. To see this we start with the second term of λ which refers to the constrained types and suppose that ϕ^{**} were to be set at its uncoordinated level ϕ^*. In this case, we know that the average distortion above the cut-off is zero so this term becomes zero. Next, we move to the first term of λ which refers to the optimising behaviour of the unconstrained types. As already discussed, the sign of g_p^* is ambiguous hence the disciplining effect is ambiguous as well. To illustrate suppose that g_p^* is positive so that higher permit prices push green investment up. In this case, a higher permit price is beneficial for countries whose type is relatively low as these types underinvest relative to their ex-ante optimum while it harms countries whose type is relatively high and overinvest relative to their ex-ante optimum. Since the disciplining effect is determined by the unconstrained low-type countries a higher permit price can increase welfare by increasing discipline which implies an increase in the cut-off value ϕ^{**}: by increasing flexibility, the coordinated rule induces a higher permit price which improves the discipline of underinvesting countries. This higher level of discipline is ex-ante beneficial for all countries.

Finally, the last term in the parenthesis in (2.19), shows the price effects on emissions and local damages and is positive (together with the minus sign). Therefore it puts a downward pressure on the threshold allowing for less flexibility and lower damages for all types.

Therefore, the optimal level of discretion in the coordinated rule and how it compares to the uncoordinated rule will depend on the magnitude of these three different effects.

2.3.2 Endogenous E

Next, we consider the optimal agreement over the cap on emissions. In reality, these decisions are taken at the same time and optimality requires that the associated first-order conditions both hold. Differentiating expected ex ante social welfare with respect to E, recognising that $\omega = E$ and taking into account (2.18) we get

$$\int_{\underline{\phi}}^{\bar{\phi}} \phi B'(y^*) f(\phi) \mathrm{d}\phi = D'(E^*) + \mathbb{E}_\phi \left[d'(e^*) \left(1 - \frac{\phi}{\phi^{**}} \beta \right) \middle| \, \phi \geq \phi^{**} \right].$$

Thus, the optimal permit allocation requires that the mean marginal benefit of energy is set equal to the global externality, plus a measure closely related to the average distortion for the constrained countries. This condition can be rewritten as

$$\frac{\mathbb{E}[\phi | \phi \geq \phi^{**}]}{\phi^{**}} = \frac{1}{\beta} + \frac{D'(E^*)}{\beta d'(e^*(\phi^{**}, p^*(\phi^{**}, E^*))} - \frac{\int_{\underline{\phi}}^{\bar{\phi}} \phi B'(y^*) f(\phi) \mathrm{d}\phi}{\beta d'(e^*(\phi^{**}, p^*(\phi^{**}, E^*))} \tag{2.22}$$

which can be combined with (2.19) to get

$$
\begin{aligned}
&\mathbb{E}\left[\phi\right] p = D'(E^*) - \frac{\mathrm{d}p^*}{\mathrm{d}E}\left(\rho(\phi^{**})\right. \\
&\left.+\lambda(\phi^{**},\beta) - \int_{\underline{\phi}}^{\phi^{**}} d'(e^*(\phi, p^*(\phi^{**}, E^*)))e_p^* f(\phi)d\phi - \int_{\phi^{**}}^{\bar{\phi}} d'(e^*(\phi^{**}, p^*(\phi^{**}, E^*)))e_p^* f(\phi)d\phi\right)
\end{aligned}
$$

where ρ and λ are as defined in Proposition 2. This can be compared with the carbon price under the constrained efficient outcome, which can be written as

$$
\mathbb{E}\left[\phi\right]\tilde{p} = D'(\tilde{E}) - \frac{\mathrm{d}\tilde{p}}{\mathrm{d}E}\left(\rho(\bar{\phi}) - \int_{\underline{\phi}}^{\bar{\phi}} d'(e^*)e_p^* f(\phi)\mathrm{d}\phi\right).
$$

That is, in the constrained efficient outcome, the planner can choose type contingent investment plans which implicitly sets the optimal cut-off at the upper bound $\bar{\phi}$ as there is no need to restrict countries further. This also implies that there is no disciplining effect and the optimal price is optimally set to take into account the marginal global and local damages and the redistributive effect which is present even in the case of no present bias, $\beta = 1$. Note that in that case, the disciplining effect becomes zero also in the coordinated action case, i.e., $\lambda(\phi^{**}, 1) = 0$ and the coordinated and constrained efficient prices can be readily compared.

Unfortunately, a general characterisation of the magnitude of the coordinated and uncoordinated investment rule is impossible without specifying functional forms for the benefit, damage and cost functions and this is our plan for the next section.

2.4 Functional forms

In this section, we use specific functional forms to get closed-form solutions and better intuition of the results. For the benefit function, we assume

$$
B(y_i) = y_i(\bar{y} - \frac{1}{2}y_i)\ , \quad B'(y_i) = \bar{y} - y_i\ , \quad B''(y_i) = -1\ , \tag{2.23}
$$

where $\bar{y} > 0$.[10] As it needs to be strictly increasing, we restrict it to the domain $y_i \in [0, \bar{y}]$. The parameter $\bar{y} \geq y_i$ denotes the ideal energy level if emitting is costless. Thus, a country would never produce more than $\bar{y}$ due to the implicit costs of generating or transporting the energy.

For the cost of green investment we assume $c(g_i) = \frac{k}{2}g_i^2$ and for damages from emissions $d(e_i) = \frac{\alpha}{2}e_i^2$ and $D(E) = \frac{\gamma}{2}E^2$.

We can then write equations (2.2) and (2.6) as

[10] The above functional form assumptions allow for analytical tractability and highlight the mechanism underlying our results. Moreover, they are common in the literature on international environmental agreements. For instance, quadratic abatement cost functions are assumed in Gersbach and Winkler (2011), Harstad (2016) and Holtsmark and Weitzman (2020).

$$\bar{y} - e_i - g_i = p \tag{2.24}$$
$$\phi_i(\bar{y} - e_i - g_i) + \beta\alpha e_i = \phi_i k g_i \tag{2.25}$$

which gives us the ex-post optimal level of emissions, investment and total energy as

$$e_i^* = \frac{\phi_i k\bar{y} - (1+k)\phi_i p}{\phi_i k + \beta\alpha} = e^*(g(\phi,p),p) = e^*(\phi,p) \tag{2.26}$$
$$g_i^* = \frac{\phi_i\beta\alpha\bar{y} + (\phi_i - \beta\alpha)p}{\phi_i k + \beta\alpha} = g^*(\phi,p) \tag{2.27}$$
$$y_i^* = e_i^* + g_i^* = \bar{y} - p = y^*(\phi,p) \tag{2.28}$$

and

$$g_\phi = -\beta\alpha(k(\bar{y}-p) - p) \leq 0 \tag{2.29}$$

since $k(\bar{y}-p) - p \geq 0$ for $e^*(\phi,p) \geq 0$

while

$$g_p = \frac{\phi_i - \beta\alpha}{\phi_i k + \beta\alpha} \lesseqgtr 0 \tag{2.30}$$

2.5 Constrained Efficiency

Combining equations (2.8) and (2.9) and taking into account that $B'(\tilde{e}^*(g^*)+g) - p = 0$ we get

$$\phi p + d'(\tilde{e}) + \alpha\tilde{e} + \gamma E - \mathbb{E}[\phi]\,p = \phi k g \qquad \text{for all } \phi \tag{2.31}$$

and solving for optimal investment g we get

$$g(\phi) = \frac{p(\phi - \mathbb{E}(\phi) - \alpha) + \gamma E + \alpha\bar{y}}{\phi k + \alpha} \tag{2.32}$$

Next taking into account $p = B'(\tilde{e}^*(g^*(\phi)) + g^*(\phi)) = \bar{y} - \tilde{e} - g^*(\phi)$ and the constraint for total emissions we get

$$E = \int_{\underline{\phi}}^{\bar{\phi}} (\bar{y} - p - g^*(\phi))f(\phi)\mathrm{d}\phi = \bar{y} - p - \int_{\underline{\phi}}^{\bar{\phi}} g^*(\phi)f(\phi)\mathrm{d}\phi$$

so that

$$\mathbb{E}(g^*(\phi)) = \int_{\underline{\phi}}^{\bar{\phi}} g^*(\phi)f(\phi)\mathrm{d}\phi = \bar{y} - p - E \tag{2.33}$$

Finally. expressing (2.32) as

$$g(\phi)(\phi k + \alpha) = p(\phi - \mathbb{E}(\phi) - \alpha) + \gamma E + \alpha\bar{y} \tag{2.34}$$

and taking expectations on both sides, we get

$$\mathbb{E}(\phi g(\phi)) = \frac{-\alpha p + \alpha \bar{y} + \gamma E - \alpha \mathbb{E}(g(\phi))}{k} \tag{2.35}$$

which taking into account (2.33) becomes

$$\mathbb{E}(\phi g(\phi)) = \frac{(\alpha+\gamma)E}{k} \tag{2.36}$$

Next, using $\frac{\mathrm{d}p^*}{\mathrm{d}E} = \left(\int_{\underline{\phi}}^{\bar{\phi}} \tilde{e}_p(\phi,p)f(\phi)\mathrm{d}\phi\right)^{-1} = B''(\tilde{e}^*(g,p)+g)^{-1} = -1$ and $p = \bar{y} - \tilde{e} - g^*(\phi)$, (2.1) becomes

$$p = \frac{(\gamma + \mathbb{E}(\phi))E + (\alpha - \mathbb{E}(\phi))\bar{y} - \alpha\mathbb{E}(g(\phi)) + \mathbb{E}(\phi g(\phi))}{\alpha} \tag{2.37}$$

Finally substituting (2.33) and (2.36), we get

$$\hat{E} = \frac{\mathbb{E}(\phi)k\bar{y}}{(1+k)(\alpha+\gamma) + k\mathbb{E}(\phi)} \tag{2.38}$$

which is the constrained efficient level of the cap on total emissions.

Finally going back to (2.32), we take expectations on both sides and making use of (2.33) and (2.38) we get

$$-p + \bar{y}\frac{(1+k)(\alpha+\gamma)}{(1+k)(\alpha+\gamma)+k\mathbb{E}(\phi)} = \mathbb{E}\left(\frac{-p(\mathbb{E}(\phi)+\alpha) + \frac{(\alpha+\gamma)(\alpha(1+k)+k\mathbb{E}(\phi))}{(1+k)(\alpha+\gamma)+k\mathbb{E}(\phi)}\bar{y} + p\phi}{\phi k + \alpha}\right) \tag{2.39}$$

which with a uniform distribution for ϕ gives us an expression for the constraint efficient level of permit price as a function of parameters only. We can then substitute this expression in (2.32) to the constrained efficient level of investment as function of parameters only.

References

Amador, M., I. Werning, and G.-M. Angeletos (2006). Commitment vs. flexibility. *Econometrica 74*(2), 365–396.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford economic papers 46*(Supplement_1), 878–894.

Bilbiie, F., T. Monacelli, and R. Perotti (2021). Fiscal policy in europe: controversies over rules, mutual insurance, and centralization. *Journal of Economic Perspectives 35*(2), 77–100.

Carraro, C. and D. Siniscalco (1993). Strategies for the international protection of the environment. *Journal of public Economics 52*(3), 309–328.

Colmer, J., R. Martin, M. Muûls, and U. J. Wagner (2024). Does pricing carbon mitigate climate change? firm-level evidence from the european union emissions trading system. *Review of Economic Studies*, rdae055.

Gersbach, H. and R. Winkler (2011). International emission permit markets with refunding. *European Economic Review 55*(6), 759–773.

Halac, M. and P. Yared (2018). Fiscal rules and discretion in a world economy. *American Economic Review 108*(8), 2305–2334.

Harstad, B. (2012). Climate contracts: A game of emissions, investments, negotiations, and renegotiations. *Review of Economic Studies 79*(4), 1527–1557.

Harstad, B. (2016). The dynamics of climate agreements. *Journal of the European Economic Association 14*(3), 719–752.

Hoel, M. (1992). International environment conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics 2*, 141–159.

Holtsmark, B. and M. L. Weitzman (2020). On the effects of linking cap-and-trade systems for CO_2 emissions. *Environmental and Resource Economics 75*, 615–630.

Leruth, B. (2024). Political long-termism and the european union: Five research questions for the future. *Contemporary European Politics 2*(2), e70000.

Martimort, D. and W. Sand-Zantman (2016). A mechanism design approach to climate-change agreements. *Journal of the European Economic Association 14*(3), 669–718.

The European Commission (2020). Commission decision (eu) 2020/2166 of 17 december 2020 on the determination of the member statesâ€™ auction shares during the period 2021-2030 of the eu emissions trading system. *Official Journal of the European Union 431*, 66–69.