Ranking suboptimal climate policies: the role of carbon markets *

Larry Karp[†]

July 24, 2020

Abstract

International negotiations have not created a foundation for a global carbon market or set emissions reduction targets that would respect a 2^oC ceiling on temperature increase: proposed abatement policies are likely both suboptimal and inefficiently implemented. The lack of a carbon market favors taxes over quotas, and the suboptimality of abatement targets favors quotas. [Now describe empical results.]

Keywords: asymmetric information, pollution control, cap and trade, Article 6 of Paris Agreement, policy ranking

JEL, classification numbers: Q000, Q500, H200, D800

---

*I thank Christian Traeger and Hiroaki Sakamoto for comments on an earlier draft. I am grateful to Max Auffammer for data suggestions and to Andy Hultgren for research assistance. The usual disclaimer applies.

†Department of Agricultural and Resource Economics, University of California, Berkeley, email: karp@berkeley.edu.

1 Introduction

The levels of carbon abatement policies recently discussed under the aegis of the United Nations Framework Convention on Climate Change (UNFCCC) are suboptimal and their proposed implementation is inefficient. The suboptimality of policy *levels* favors quantity restrictions over taxes, but the lack of a carbon market favors taxes. I calculate these offsetting effects and then estimate their magnitudes. For a given level of aggregate emissions, the efficient cross-country allocation of emissions permits based on observables, but without international trade in permits, would achieve approximately xx% of the potential gains.

Current national pledges under the 2016 Paris Agreement, if honored, would achieve less than a third of the carbon emissions reductions needed to keep global temperatures from rising above 2^oC (Piris-Cabezas et al. 2018). By this measure, proposed policy targets are insufficiently stringent. Mehling et al. 2018 and Edmonds et al. 2019 estimate that international trade in carbon permits could reduce the cost of achieving the Paris Agreement pledges by $30\% - 75\%$. That cost reduction would make it easier for countries to adopt the stricter carbon limits needed meet the 2^oC target. Some of this potential cost savings could be achieved by the allocation of carbon permits based on observable (verifiable) information, but markets – or a more complicated mechanism for revealing information – are needed to achieve full efficiency. Current estimates do not break down the potential gains arising from observable versus nonverifiable information.

COP 25, the 2019 Madrid meetings of the UNFCCC, sought to flesh out the details of Article 6.2 of the Paris Agreement. This article was designed to lay the foundation for a global carbon market, and more generally to enable countries to exchange credits for reductions achieved by different types of policies (e.g. cap and trade and a carbon tax). COP 25's failure to achieve this objective illustrates the political difficulty of establishing an international carbon market – and the still greater difficulty of establishing rules for transfers across heterogenous policies.

Weitzman's (1974) classic "Prices Versus Quantities" ranks price-based and quantity-based policies. His paper and subsequent contributions recognize that the ranking criterion applies to taxes and cap and trade, not to taxes and *untraded* quotas (Stavins 2019). Both taxes and cap and trade result in the efficient allocation of emissions across agents with different abatement costs. In contrast, with untraded quotas, different firms or regions typically have different equilibrium levels of marginal cost, resulting in inefficiency.

Several papers generalize Weitzman's static model to a dynamic setting, making it appropriate for stock pollutants such as greenhouse gasses: Hoel and Karp (2002) compare policy ranking under the open loop versus feedback settings; Newell and Pizer (2003) study the role of serially correlated shocks in the open loop setting. Karp and Traeger (2019) introduce gradual diffusion of technology shocks and also explain why the feedback policy ranking depends on the intercept in addition to the slope of the social cost of carbon (defined as the additional present value of future costs arising from an additional unit of current emissions.) These papers, like Weitzman's, assume that under the cap there is trade in emissions permits, and they compare the information-constrained *optimal* tax and cap and trade policies.

Current policy proposals are far from optimal, and international carbon markets are rudimentary. Existing markets link the European Union (EU) members, California with Quebec, and the EU with Switzerland. How should we rank price-based and quantity-based policies without international trade in carbon, and/or when the policies under consideration are suboptimal?

I begin with a variant of Weitzman's static model. Although usually discussed in a representative agent setting, Section 5 of his paper uses a multi-agent model to show that the ranking criterion involves cap and trade, not cap without trade. With slightly more structure, the multi-agent economy produces an intuitive measure of the gains from trade – by which I always mean the gains in excess of those achieved by allocating permits efficiently based on observable heterogeneity. The ranking criterion without trade depends on the variances of the idiosyncratic and systemic shock and (as usual) on the ratio of slopes of marginal damages and marginal abatement cost. With positive gains from trade, the lack of a market favors taxes.

For both a flow and a stock pollutant, the per-period gains from trade are invariant to policy stringency: they apply to both optimal and suboptimal policy pairs that lead to the same *expected* level of emissions ("certainty equivalent" policies). This invariance makes it easy to evaluate the importance of markets for emissions permits under non-optimal policies. The gains from trade depend only on the variance of the idiosyncratic shocks and on the slope of the marginal abatement cost. These are exogenous parameters, and thus are independent of policy stringency in both the flow and the stock pollutant settings. This exogeneity explains the invariance for both settings.

For a flow pollutant, the policy ranking both with and without trade in permits is also invariant to the stringency of policy. The ranking criterion under trade depends only on the ratio of slopes of marginal abatement costs and marginal damages; the ranking criterion absent trade additionally depends on the variances of systemic and idiosyncratic shocks. These are all exogenous parameters for flow pollutants, explaining the invariance result.

In contrast, with a stock pollutant where a regulator conditions policies on current information, the policy ranking with or without trade depends on the stringency of policy.[1] With convex stock-dependent damages, the marginal welfare cost of an additional unit of emissions depends on *future* policies. If future policies are lax, then future emissions will be high, resulting in a relatively high and steep social cost of carbon. Stringent future policies lead to a lower and flatter social cost of carbon. Because policy ranking (with or without trade) depends on the endogenous social cost of carbon, it is not invariant to the stringency of certainty equivalent policies. By treating the actual policy as a convex combination of the optimal policy and unregulated emissions, I obtain a one-parameter family of policy rules. A laxer policy implies a higher and steeper social cost of carbon, favoring the quantity-based policy.

Thus, actual policy limitations – suboptimal levels and lack of markets – have offsetting effects on policy ranking. The lack of international markets for carbon permits favors taxes, and the suboptimality of proposed policies favors quotas. The formulae I provide, together with parameter estimates taken from the literature and new econometric estimates, suggest that xxxx.

Article 6.2 of the Paris Agreement seeks to create a framework for the international exchange of reductions achieved by different policies, not only for the establishment of international carbon markets. Price and quantity policies are of course not strictly equivalent – if they were, there would be no scope for welfare ranking. However, for every quota there is a tax that produces the same expected level of emissions. These certainty equivalent policies therefore provide a good setting for thinking about Article 6.2.

There are at least two ways to create a many-agent model that generalizes the familiar representative agent model: the accepted method that Weitzman and others use, and which I adapt, and an alternative that merely adds

[1] Newell and Pizer (2003), footnote 14, note that the policy ranking for a stock pollutant does not depend on policy stringency in the open loop setting. There the regulator chooses the sequence of future policies at time zero, so information is static. Thus, the dependency of the ranking on policy stringency arises from the assumption that future policies are conditioned on information that becomes available in the future – not on the change from a flow pollutant to a stock pollutant.

agents to the representative agent model. The alternative alters the feasible set, thereby conflating changes in outcome arising from changes in strategic incentives (with many agents) and changes due to the altered feasible set. I discuss this difference when setting up the model, and then illustrate its importance in a final section.

Gains from trade Copeland and Taylor (2005) note that when the conditions for Factor Price Equalization approximately hold, the lack of an international market for emissions permits causes only a small loss in welfare. Arkolakis, Costinot, and Rodriguez-Clare (2012) estimate U.S. gains from trade at 1.4% of 2000 GDP. Cosinot and Rodriques-Clare's (2018) review reports estimates of the gains for trade from a large economy such as the US at 2- 8% of GDP. Fally and Sayre (2019) estimate the gains from trade in a model that disaggregates across agricultural and natural resource commodities. These commodities' low elasticities of supply and of substitution lead to much larger estimates of gains from trade: twice as large as previous estimates for large countries, three times as large for medium-sized countries, and many times as large for small countries.

Rafey (2019) examines the effect of creating a domestic water market in Australia. He estimates that this market increased the volume of irrigated agricultural output in Australia by 4- 6%, producing gains from trade of 8 - 12% of the value of irrigated agricultural production.

All of these estimates are much smaller than the gains arising from trade in emissions permits, consistent with the 30- 70% reduction in abatement cost reported above.

2 Weitzman's model and ranking criterion

A tax and quota pair are certainty equivalents if and only if the expected emissions under the tax equals the quota. I regard a certainty equivalent tax and quota combination as "admissible" if and only if, with a "sufficiently high probability" (as determined by the modeler) the quota is binding and the tax induces positive emissions.

A very lax quota would be slack with high probability, and a very stringent tax would induce zero emissions with high probability. In either of those cases, the propositions below would rest on invalid assumptions. Thus, although the tax and quota combinations described in subsequent propositions might be non-optimal, they cannot be extremely lax or extremely restrictive in relation to the distribution of the shock. Appendix C.1 provides a precise definition of "admissibility" and shows that under reasonable circumstances a broad range of policies are admissible with probability 1. Weakening the requirement by reducing the probability enlarges the set of admissible policies.[2] Hereafter I assume that all policies under consideration are admissible.

Weitzman's slope-based ranking criterion applies to any certainty equivalent policy pair, including the optimal pair. The firm in the Representative Agent Economy (RAE) obtains the private benefit $(B_0 + \theta) E - \frac{B}{2} E^2$ of emitting E; the random variable (private information) is $\theta \sim (0, \sigma^2)$.[3] Society incurs the pollution cost (external to the polluter) $D_0 E + \frac{D}{2} E^2$.Society's payoff is quadratic in E and linear in θ:

$$\underbrace{(B_0 + \theta) E - \frac{B}{2} E^2}_{\text{global private benefit of emissions}} \quad - \underbrace{\left(D_0 E + \frac{D}{2} E^2 \right)}_{\text{global damage of emission}} . \tag{1}$$

Facing a tax, τ, the representative agent maximizes $(B_0 + \theta) E - \frac{B}{2} E^2 - \tau E$, resulting in emissions $E(\tau) = \frac{B_0 - \tau + \theta}{B}$. The expectation and the variance of emissions given the tax are, respectively,

$$\bar{E}(\tau) = \frac{B_0 - \tau}{B} \text{ and } \frac{\sigma^2}{B^2}. \tag{2}$$

The representative agent observes θ, but the planner knows only its distribution. The planner can control emissions using either a tax, τ, or a quota, Q. The regulator who uses the quota Q obtains the expected payoff

$$\begin{aligned} W^Q(Q) &\equiv \mathbf{E}_\theta \left[(B_0 + \theta) Q - \tfrac{B}{2} Q^2 - \left(D_0 Q + \tfrac{D}{2} Q^2 \right) \right] \\ &= \underline{B_0 Q - \tfrac{B}{2} Q^2 - \left(D_0 Q + \tfrac{D}{2} Q^2 \right)}. \end{aligned} \tag{3}$$

---

[2] With high probability, the optimal tax might generate zero emissions or the optimal quota might be slack . Most of the literature following Weitzman (1974) ignores this issue.

[3] Treating actions as emissions, a public bad, or abatement, a public good, are equivalent, but the former corresponds directly to the model estimated in Section 6. The random variable θ corresponds to a demand shock for emissions or an abatement cost shock.

The confirm the equivalence of treating the action as emissions or abatement, use the fact that the unregulated level of emissions is $E^{BAU} = \frac{B_0 + \theta}{B}$. When emissions are restricted to E, define abatement as $A(E, \theta) \equiv E^{BAU} - E$, so $\frac{\partial A}{\partial E} = -1$. Define the cost of abatement, $C(A(E, \theta))$, as the reduction in benefit due to the reduction in emissions: $C(A(E, \theta)) = \frac{1}{2} \frac{(B_0 + \theta)^2}{B} - \left((B_0 + \theta) E - \frac{B}{2} E^2 \right)$. The marginal cost of abatement, $\frac{\partial C}{\partial A} = \frac{\partial C}{\partial E} \frac{dE}{dA} = B_0 + \theta - BE$, equals the marginal benefit of emissions.

The quota enables the regulator to choose the actual level of emissions.

If the regulator uses the tax τ, emissions are stochastic, equal to $\bar{E}+\frac{\theta}{B}$. By modeling the tax-setting regulator as choosing the expected level of emissions, $\bar{E}(\tau)=\frac{B_0-\tau}{B}$ (instead of the tax, τ) the payoff becomes

$$\begin{aligned} W^{\tau}\left(\bar{E}\right) \equiv \mathbf{E}_{\theta}[&(B_0+\theta)\left(\bar{E}+\tfrac{\theta}{B}\right)-\tfrac{B}{2}\left(\bar{E}+\tfrac{\theta}{B}\right)^2 \\ &-\left(D_0\left(\bar{E}+\tfrac{\theta}{B}\right)+\tfrac{D}{2}\left(\bar{E}+\tfrac{\theta}{B}\right)^2\right)]. \end{aligned} \tag{4}$$

Taking expectations, using $\mathbf{E}\theta=0$ and $\mathbf{E}\theta^2=\sigma^2$ gives

$$W^{\tau}\left(\bar{E}\right)=\underline{B_0\bar{E}-\frac{B}{2}\bar{E}^2-\left(D_0\bar{E}+\frac{D}{2}\bar{E}^2\right)}+\frac{1}{2B}\left[1-\frac{D}{B}\right]\sigma^2. \tag{5}$$

Inspection of equations 3 and 5 shows that payoffs are additively separable in the policy (either Q or $\bar{E}$) and the variance, σ^2. Therefore, the optimal action is independent of σ^2 in both cases. Consequently, the optimal quota and the optimal tax under uncertainty equal their respective levels under certainty. This result, which follows from the quadratic structure with additive uncertainty, is known as the Principle of Certainty Equivalence.

Because the underlined terms in equations 3 and 5 are the same, apart from the names given their arguments, the maximizing controls, Q^* and $\bar{E}^*$, are equal: the optimal policies are certainty equivalents. Moreover, for *any* certainty equivalent policy pair, $Q=\bar{E}$

$$W^{\tau}\left(\bar{E}\right)-W^{Q}\left(Q\right)=\frac{1}{2B}\left[1-\frac{D}{B}\right]\sigma^2. \tag{6}$$

Taxes welfare-dominate quotas if and only if $B>D$, producing Weitzman's ranking criterion. It applies to any certainty equivalent policy pair.

Moving from a quota to its certainty equivalent tax creates a benefit and a cost to society. The benefit, due to the representative firm's ability to respond to shocks, equals $\frac{1}{2B}\sigma^2$. However, emissions are stochastic under the tax, but deterministic under the (binding) quota. Because the policies produce the same expected level of emissions, Jensen's inequality implies that expected damages are higher under the tax. The increase in expected damages under the tax equals $\frac{1}{2B}\frac{D}{B}\sigma^2$. Therefore, equation 6 gives the net benefit of moving from the quota to its certainty equivalent tax.

3 A model with many agents

There are at least two ways of creating a many-agent economy related to the representative agent model above ("RAE"). Following Weitzman, I use a "fragmented economy" (FE), one that splits the RAE into n agents, but *without changing the technology* and thus without changing the feasible set. The alternative, discussed below, simply adds agent to the one-agent RAE.

Ex post heterogenous agents receive different shocks; these are private information, or at least not verifiable. Ex ante heterogenous agents also have publicly observed differences, known to the regulator. Ex ante heterogeneity is essential when taking the model to data and also in evaluating the practical importance of an international carbon market.

Quota shares are "information-constrained efficient" if and only if they equalize firms' expected marginal benefit of emissions, conditional on the publicly observed information. Absent trade in permits, a regulator achieves information-constrained efficiency by an appropriate distribution of shares. For example, if agents have no observable differences (ex ante homogeneity), then the information-constrained efficient allocation gives each agent an equal share of the aggregate quota. With asymmetric information (ex post heterogeneity), full efficiency requires trade in permits, or a more complicated mechanism to reveal private information.

Comparability of the RAE and the FE requires that they produce the same aggregate results, given the same tax or the same aggregate quota. That is, they must be observationally equivalent. More formally:

Definition 1 *Observational equivalence holds if and only if for all integers $n \geq 1$: (i) If the policy is a quota, assume that in the FE the quota is not tradeable but the allocation of shares is information-constrained efficient. For every such quota E, the expected aggregate payoff in the FE equals the expected payoff in the RAE.*[4] *(ii) For every tax, the mean and variance of aggregate emissions are the same in the RAE and the FE.*

Fragmentation makes it easy to examine the effects of introducing trade in permits in an economy constrained by an aggregate cap on emissions.

[4] I could replace Definition 1.i with the requirement that the expected payoff in the FE *with* trade equals the expected payoff in the RAE, given the same aggregate quota. This alternative changes nothing of substance because it does not change the measure of the gains from trade.

Increasing n by fragmenting the economy differs from simply adding another agent to an existing economy. Adding another agent makes the economy larger, thereby increasing the number of agents who create and then suffer from the public bad or enjoy the public good. In contrast, fragmenting the economy creates no *intrinsic* changes. It might change agents' strategic incentives, thereby changing the outcome, but it does not change the feasible set; increasing n by adding another agent to the economy changes the feasible set. If Brexit causes the UK and the EU climate policies to diverge, the number of non-cooperative countries increases, changing strategic incentives. However, the environmental problem does not change unless some country changes their behavior. Section 7 provides an example of the different implications of the two ways of increasing n.

The plausible magnitude of n depends on the context. To examine the importance of linking regional carbon markets (e.g. in California, a group of Northeastern states and some Canadian provinces, or across regions in China) or for comparing noncooperative behavior amongst blocs of countries, n is likely small. To examine the importance of inter-firm trade within a nation, n may be in the hundreds, the number of firms subject to the cap. Letting $n \to \infty$ in the fragmented model leads to a sensible representation of an economy with a continuum of heterogenous firms. Letting $n \to \infty$ by adding agents means that marginal environmental damages become unbounded, driving the socially optimal level of pollution to zero; this result is not a sensible representation of an economy with many heterogenous firms.

The benefit to agent i in the FE of emitting e_i is $a + (b_{0i} + \theta_i) e_i - \frac{b}{2} e_i^2$.[5] Define a measure of ex ante (i.e., observable) heterogeneity:

$$\widehat{var\left(b_{0i}\right)} \equiv \left(\frac{1}{n}\sum_i b_{0i}^2\right) - \left(\frac{1}{n}\sum_i b_{0i}\right)^2. \tag{7}$$

Agent i's shock is $\theta_i = \alpha + \mu_i$; α is the systemic shock common to all agents, and μ_i is the idiosyncratic (agent-specific) shock, with[6,7]

$$\alpha \sim \left(0, \sigma_\alpha^2\right), \ \mu_i \sim iid\left(0, \sigma_\mu^2\right), \text{ and } \mathbf{E}\left(\alpha\mu_i\right) = 0. \tag{8}$$

[5] I normalize the constant in the RAE payoff to zero. Payoffs in the RAE and the FE are both quadratic functions of emissions, but their coefficients differ.

[6] Due to the quadratic structure, results depend on only the first two moments of the distribution, 0 and σ^2, not on on the form of the distribution, e.g. the higher moments. Those features do affect the admissibility requirement; see Appendix C.1.

[7] The assumption $\mu_i \sim iid\left(0, \sigma_\mu^2(n)\right)$ implies that agents' shocks, θ_i, are positively correlated; however conditional on the systemic shock α, agents' shocks are uncorrelated.

Agents are ex post heterogenous if and only if $\sigma_\mu^2 > 0$. Facing a tax τ, agent i in the FE chooses emissions to maximize $(b_{0i} + \theta_i) e_i - \frac{b}{2} e_i^2 - \tau e_i$, resulting in emissions $e_i(\tau) = \frac{b_{0i} - \tau + \theta_i}{b}$.

Lemma 1 *(i) Observational equivalence, Definition 1, holds if and only if*

$$\sum_i b_{0i} = nB_0;\ b = nB;\ a = -\frac{\widehat{var(b_{0i})}}{2nB} \text{ and} \tag{9}$$

$$\sigma^2 = \sigma_\alpha^2(n) + \frac{\sigma_\mu^2(n)}{n}. \tag{10}$$

(ii) In the absence of trade in emissions quotas, the increase in expected welfare due to using information-constrained optimal quota shares, rather than simply giving each agent an equal share, is $-na = \frac{1}{2B}\widehat{var(b_{0i})}$.

Lemma 1 (i) is trivial if agents are ex ante homogenous, where $\widehat{var(b_{0i})} = 0$. There, I obtain the first two parts of equation 9 by equating aggregate payoffs in the RAE and the FE under an arbitrary aggregate quota, E:

$$\begin{gathered}\mathbf{E}_\theta (B_0 + \theta) E - \tfrac{B}{2} E^2 - \left(D_0 E + \tfrac{D}{2} E^2\right) = \\ n\mathbf{E}_{\{\theta_i\}} \left[a + (b_0 + \theta_i) \tfrac{E}{n} - \tfrac{b}{2}\left(\tfrac{E}{n}\right)^2\right] - \left(D_0 E + \tfrac{D}{2} E^2\right).\end{gathered}$$

Equating coefficients produces the first two parts of equation 9. If agents are ex ante heterogenous and receive information-constrained optimal quotas, agents with larger b_{0i} have a larger marginal benefit of emissions and receive a larger quota. By Jensen's Inequality, the expected benefit of an aggregate level of emissions is greater, the larger is the dispersion of the b_{0i}'s. Therefore, the constant term in each agent's payoff, a, must be negative in order that expected aggregate benefits are the same in the FE and the RAE. Equation 10 follows from the requirement that the variance of aggregate emissions is the same in the FE and the RAE for all taxes (Definition 1.ii).

To emphasize that there is a continuum of solutions to equation 10, I write $\sigma_\alpha^2(n)$ and $\sigma_\mu^2(n)$ as functions; for any non-negative function $\sigma_\alpha^2(n) \le \sigma^2$, there is a non-negative function $\sigma_\mu^2(n)$ satisfying equation 10.[8]

[8] The first two equations in system 9 reproduce Weitzman's equations 24 and 25. Weitzman does not include ex ante heterogeneity, and therefore does not consider the constant

4 Policy ranking in the Fragmented Economy

Section 2 notes the restriction to "admissible" policies, those for which the quota is binding and the tax induces positive emissions (with sufficiently high probability, as determined by the modeler). In the fragmented economy I also require that *all* agents emit at positive levels under both the tax and under the tradeable quota. This additional requirement is very weak, particularly if the agents are countries or regions. The requirement means that the variance of idiosyncratic shocks cannot be arbitrarily large.

Observational equivalence means that an arbitrary tax results in the same mean and variance of emissions in the RAE and in the FE. The damage component of welfare, $\mathbf{E}\left(D_0 E + \frac{D}{2}E^2\right)$, depends only on the mean and variance of aggregate emissions. Moreover, the Principle of Certainty Equivalence implies that the optimal policy levels do not depend on second or higher moments of the shock. Therefore, the optimal tax, τ^*, is the same in the RAE and in the FE, as is the optimal quota (with or without trade), Q^*.

Proposition 1 *Assume that, absent trade, the allocation of quota shares in the FE is information-constrained efficient. For any admissible certainty equivalent tax and quota pair, including the optimal levels:*

(i) The gains from trade under a quota are

$$G = \frac{n-1}{2nB}\sigma_\mu^2. \tag{11}$$

(ii) Taxes dominate the quota without trade if and only if

$$k + 1 - \frac{D}{B} > 0 \text{ with } k \equiv \frac{(n-1)\,\sigma_\mu^2}{n\sigma^2}. \tag{12}$$

(iii) The welfare ranking is the same in the RAE and in the FE with $\sigma_\mu^2 > 0$ *if and only if agents can trade emissions permits.*

(iv) (a) The optimal tax, τ^**, is the same in the RAE and in the FE. (b) The optimal quota,* Q^**, is the same in the RAE and in the FE with or without trade. (c) Welfare under the optimal quota is the same in the RAE and in*

a. As noted in the text, ex ante heterogeneity is needed for the empirical application and also to assess the welfare gain arising from markets, over an above the gains achieved by ex ante efficient allocation of quotas. The degree of freedom in equation 10 parallels the indeterminacy beneath Weitzman's equation 29. Equation 10 produces a simple parameterization of this indeterminacy.

the FE without trade. (d) Under the optimal tax, the difference in welfare between the FE and the RAE equals the gains from trade.

Parts (i) and (ii) are new.[9] Part (i) shows how the gains from trade depend on n, B, and σ_μ^2. Part (ii) shows how the ranking criterion changes without trade in emissions permits. The critical ratio of slopes $\frac{D}{B}$ below which taxes dominate, increases from 1 to $1+k$ without trade.[10]

Both the gains from trade and the ranking criterion are invariant to the certainty equivalent policy pair. This result, which echoes the invariance noted below equation 6, is significant because often we want to compare two non-optimal policies that achieve the same expected result.

The function $\widehat{var\left(b_{0i}\right)}$ measures observable differences across agents, and the variance of the idiosyncratic shock, σ_μ^2, measures unobserved differences. Absent trade, the distribution of permits that takes into account observed differences in the demand for emissions (instead of giving each agent an equal quota) increases welfare by $\frac{\widehat{var(b_{0i})}}{2nB}$ (Lemma 1.ii). Allowing trade in permits achieves the additional gain G. The relative importance of these two sources of heterogeneity is

$$r \equiv \frac{G}{\frac{\widehat{var(b_{0i})}}{2nB}} = \frac{\left(n-1\right)\sigma_\mu^2}{\widehat{var\left(b_{0i}\right)}}. \tag{13}$$

With this definition, the efficient distribution of permits, based on observable differences, achieves $\frac{r}{1+r}100\%$ of the gain that could be achieved with an efficient market for permits.[11]

[9] Part (iii) reproduces Weitzman's (1974) widely known result on the importance of trade. This result is due to the fact that the welfare increase under taxes, in moving from the RAE to the FE, equals the gains from trade under the quota. Part iv collects implications of the Principle of Certainty Equivalence, and is used in proving Parts i–iii.

[10] Suppose that we hold σ^2 fixed and change n, i.e. we fragment the economy to different degrees. From equation 10, σ_α^2 and/or σ_μ^2 must vary with n. From equations 11 and 12, in the limit as $n \to \infty$, $k \to \frac{\sigma_\mu^2(\infty)}{\sigma^2}$ and $G \to \frac{1}{2B}\sigma_\mu^2(\infty)$. Holding both σ^2 and n fixed, equation 10 implies $0 \le \sigma_\mu^2 \le n\sigma^2$, so $0 \le k \le n-1$.

[11] Referees' Appendix C.3 provides one additional result, showing that an agent's emissions are negatively correlated with the emissions price under cap and trade if and only if $n > 2$ and $\sigma_\mu > 0$.

5 A stock pollutant

The preceding analysis concerns a flow pollutant, where damages arise from contemporaneous emissions. This section studies the dynamic setting where damages arise instead from a stock pollutant such as greenhouse gasses. The difference between the current pollution stock and the damage-minimizing level (e.g. the preindustrial level in the case of greenhouse gasses) is S_t. This stock obeys the difference equation $S_{t+1} = \delta S_t + E_t$; the parameter $0 < \delta \leq 1$ measures the persistence of the stock.

To introduce persistence of shocks in the RAE, previous papers set $\nu_t = \rho\nu_{t-1} + \theta_t$, where ν_{t-1} is public knowledge at t, and $\theta_t \sim iid\,(0, \sigma^2)$ is the aggregation of firms' private information at t. The marginal benefit of emissions equals $B_{0t} + \nu_t - BE_t$, where B_{0t} possibly includes both a time trend and exogenous variables that affect the demand and/or supply of emissions.[12] Previous papers assume that there is within-period trade in emissions permits; in addition, like the earlier literature that studies flow pollutants, they compare welfare under the optimal tax and the optimal cap-and-trade policy.

Greenhouse gasses are a global pollutant. Reducing emissions requires international cooperation, and an efficient market requires international trade in permits. It is politically harder to establish a global compared to a national market for carbon. In addition, proposed abatement levels are suboptimal relative to the 2^oC target. How does the lack of an international market for carbon and/or the non-optimality of policies affect the ranking of price- and quantity-based policies? To answer this question, I first modify the standard model to account for non-optimal policies. I then fragment this model in order to examine the importance of trade in emissions permits.

The state variable is the triple (t, S_t, ν_{t-1}). The optimal tax and quota at t, with or without trade in permits, are certainty equivalents: they produce the same expected level of emissions, a linear function of (S_t, ν_{t-1}). The unregulated expected level of emissions is $E_t^{BAU} = \frac{B_{0t}+\rho\nu_{t-1}}{B}$, which is also (trivially) linear in (S_t, ν_{t-1}). For $c \in [0, 1]$, denote $E_{t+\tau}^c\,(t, S_t, \nu_{t-1}; c)$ as a convex combination of the optimal ($c = 0$) and the unregulated ($c = 1$) expected emissions levels. I use c as a shorthand for policy. Actual emissions

[12] The regulator learns ν_{t-1} by observing the aggregate response to the previous tax, or the quota price under the previous quota. In the RAE trade under quotas is zero, but the price that supports this equilibrium is monotonic in the lagged shock. The time dependence of $B_{0,t}$ is important when estimating or calibrating the model, but it does not change the ranking criteria – another consequence of the Principle of Certainty Equivalence.

at $t+\tau$, $\tau \geq 0$ under policy c equal

$$E_{t+\tau}^{c,i} = \begin{cases} E_{t+\tau}^{c}\left(t, S_t, \nu_{t-1}; c\right) \text{ for } i = \text{quota} \\ E_{t+\tau}^{c}\left(t, S_t, \nu_{t-1}; c\right) + \frac{\theta_{t+\tau}}{B} \text{ for } i = \text{tax} \end{cases}. \tag{14}$$

Emissions under the tax equal emissions under the quota, plus a random term that arises from the firm's response to the tax and the shock.

Economic models (e.g. Hassler et al. 2017) have begun to use the recent evidence suggesting a near-linear relation between cumulative CO_2 emissions and changes in global temperature (MacDougall et al 2017, Mathews et al. 2009). Using this linearity hypothesis, and the estimate that current pledges under the Paris Agreement would achieve about one third of the reductions needed to satisfy the (assumed optimal) 2^oC target, then these pledges correspond to $c = 0.67$.

With discount factor β, the expected payoff in the RAE under policy c, i, with $c \in [0, 1]$ and $i \in \{\text{tax,quota}\}$ is

$$J^{c,i}\left(c; t, S_t, \nu_{t-1}\right) \equiv$$
$$\underset{\{\theta_{t+\tau}\}_{\tau=0}^{\infty}}{\mathbf{E}} \sum_{\tau=0}^{\infty} \beta^{\tau} \left[\left(B_{0t+\tau} + \theta_{t+\tau}\right) E_{t+\tau} - \frac{B}{2} E_{t+\tau}^2 - \left(D_0 S_{t+\tau} + \frac{D}{2} S_{t+\tau}^2\right)\right], \tag{15}$$

where the state variables (S_t, ν_{t-1}) obey the difference equations given above and emissions are given by equation 14. The flow payoff in this model (the function in square brackets) is the same as in the static model, except that damages here depend on the pollutant stock, S, not the flow, E. Denote $\Delta \equiv J^{c,tax}\left(c; t, S_t, \nu_{t-1}\right) - J^{c,quota}\left(c; t, S_t, \nu_{t-1}\right)$, the expected advantage of taxes rather than quotas under policy c in the RAE.

Lemma 2 *The expected payoff advantage of taxes rather than quotas in the RAE with policy c is the constant $\Delta = g\left(c; D, B, \rho, \beta, \delta\right) \frac{\sigma^2}{2(1-\beta)B}$. Taxes dominate quotas iff $g\left(c; D, B, \rho, \beta, \delta\right) > 0$.*

Lemma 2 generalizes Proposition 1 of Karp and Traeger (2019), which considers the case of optimal policies ($c = 0$).[13] The function $g\left(0; D, B, \rho, \beta, \delta\right)$, corresponding to optimal policy, is simple enough to produce comparative statics, but for $c > 0$ I require simulations. Therefore, I relegate the formula for $g\left(c; D, B, \rho, \beta, \delta\right)$ to Appendix C.2.

[13] Karp and Traeger (2019) model gradual diffusion of the shock. Here I consider the special case of immediate diffusion, but the proof of Lemma 2 includes gradual diffusion.

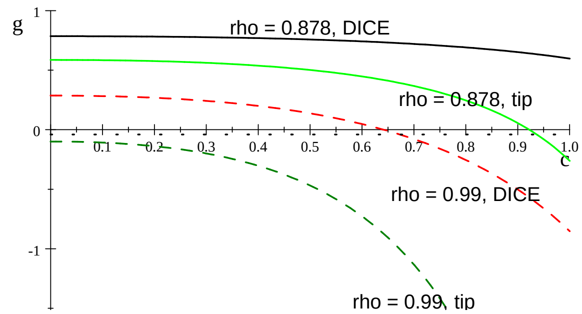

Figure 1: Taxes dominate quotas if and only if g>0. An increase in c corresponds to a less stringent policy; $c = 0$ corresponds to the optimal policy. Curves labelled "DICE" correspond to $D = 0.0015$, calibrated from DICE. Curves labelled "tip" (for "tipping") correspond to $D = 0.0045$, a value consistent with Weitzman (2012). Dashed curves correspond to $\rho = 0.878$, the point estimate from Section 6. Solid curves correspond to $\rho = 0.99$, equal to the value used in Karp and Traeger (2019) 6. Other parameter values taken from Table 1 of Karp and Traeger (2019), with an annual time step and β=0.98, δ=0.997, and B=1.8456.

Figure 1 shows the graphs of g using the point estimate $\rho = 0.878$ from Appendix B, and the value $\rho = 0.99$ taken from Karp and Traeger (2019); that value is within the 95% confidence interval of the point estimate below. The curves labelled "DICE" corresponds $D = 0.0015$, a value obtained using a quadratic approximation of DICE 13 (Nordhaus 2013); the curve "tip" (for "tipping") corresponds to steeper marginal damages, along the lines suggested by Weitzman (2012).[14] The remaining parameters, $B = 1.8456$, $\beta = 0.98$ and $\delta = 0.997$ are taken from Table 1 in (Karp and Traeger 2019).

The new result is that as regulation becomes weaker, i.e. as the policy moves away from optimality (c increases), the relative advantage of taxes also weakens. A less stringent policy increases current and future emissions, leading to higher future pollution stocks, increasing the slope and intercept of the social cost of carbon (or more generally, of the shadow cost of the pollutant). These changes favor quotas.

[14] The "DICE" scenario is calibrated assuming that an increase of atmospheric stock of 1270 GtCO2 causes a medium to long run temperature increase of $2^o C$, and that this temperature change rduces output by 1%. The "tipping" scenario assumes that this temperture chage reduces output by 3%. Thus, even the tipping scenario assumes that damages are moderate. The units of both B and D are $\frac{G\$}{(GtCO2)^2}$.

I use the fragmented economy to examine the importance of trade in carbon permits. The dynamic analog of the static FE imposes the constraints in equations 9 and 10 and sets country i's shock in period t to $\nu_{i,t}$, with

$$\nu_{i,t} = \nu_t + \mu_{i,t} - \frac{\sum_j \mu_{jt}}{n}, \text{ with } \nu_t = \rho\nu_{t-1} + \alpha_t + \frac{\sum_j \mu_{j,t}}{n} \Rightarrow$$

$$\nu_{i,t} = \rho\nu_{t-1} + \alpha_t + \mu_{i,t}, \text{ with} \tag{16}$$

$$\alpha_t \sim iid\left(0, \sigma_\alpha^2\right), \mu_{i,t} \sim iid\left(0, \sigma_\mu^2\right), \mathbf{E}\left(\alpha_t \mu_{i,\tau}\right) = 0 \forall t, \tau.$$

This model satisfies observational equivalence and it accommodates intertemporal correlation of aggregate shocks. The innovation in the RAE, θ_t, equals the aggregate innovation in the FE:

$$\theta_t = \alpha_t + \frac{\sum_j \mu_{jt}}{n}, \text{ with } var\left(\theta_t\right) = \sigma^2 = \sigma_\alpha^2 + \frac{\sigma_\mu^2}{n},$$

reproducing equation 10.

By the Principle of Certainty Equivalence, the optimal decision rule for the quota does not depend on whether there is trade in permits, just as in the static setting. Moreover, the single period gains from trade, G, are invariant to the quota. Therefore, the present discounted expected value of the gains from trade equals $\frac{G}{1-\beta}$. As Figure 1 illustrates, the ranking criterion g depends on the stringency of policy.

Proposition 2 *(a) Absent carbon markets and given policy c, taxes dominate quotas if and only if $g(c; D, B, \rho, \beta, \delta) + k > 0$, i.e. if and only if $\frac{-g}{k} < 1$. (b) With carbon markets (or for $\sigma_\mu^2 = 0$), taxes dominate quotas if and only if $g(c; D, B, \rho, \beta, \delta) > 0$.*

6 Estimation

This section estimates k, the correction needed to account for the lack of an international market for carbon permits, and r, the measure of the relative importance of unobserved and observed heterogeneity. I use annual data of carbon emissions (source) from 1945 to 2005 and divide the world into $n = 4$ regions: the US (United States), the EU (European Union), BRIC (Brazil, Russia, India and China) and Other (the Rest of World). During this period, observed emissions correspond to Business as Usual (BAU) emissions.

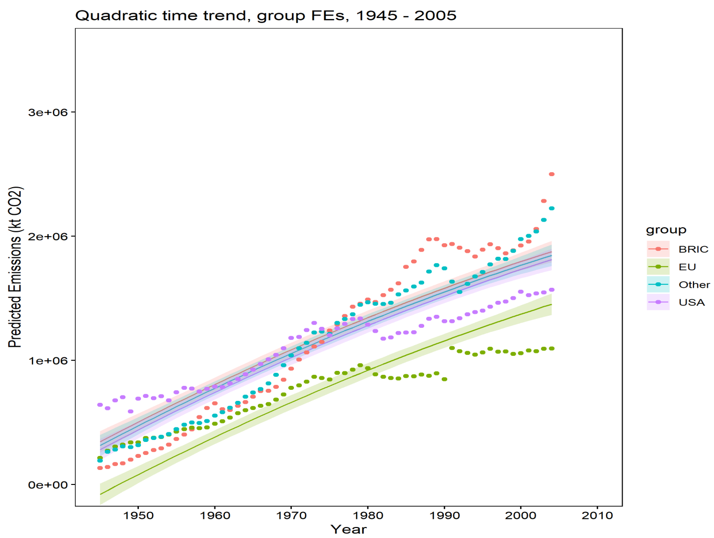

Figure 2: Carbon emissions in the four regions during 1945 - 2005

Figure 2 shows the time series of carbon emissions in the four regions over 1945 – 2005, together with curves constructed using region-specific constants and a common quadratic time trend, $h(t) = c_1 t + c_2 t^2$. Region i's marginal benefit of emissions is $b_{0i} + h(t) + \nu_{it} - be_{it}$, so absent regulation we observe its BAU emissions

$$e_{it} = \frac{1}{b}\left(b_{0i} + h(t) + \nu_{i,t}\right), \tag{17}$$

The fixed effect b_{0i} accounts for observed regional heterogeneity, and ν_{it} is region i's shock in period t.

Appendix B explains why I do not use a longer data set and also discusses a generalization that replaces the common time trend with a region-specific function $h_i(t, x_{i,t})$. If $x_{i,t}$ includes data on regulations, it possible to use the model with data on regulated pollutants. The key assumption is that the slope parameter b is constant over regions and over time. The distributional assumptions in equation 16 could be relaxed, but the current model produces a simple expression for the elements of the covariance matrix corresponding to equation 17:[15]

$$\frac{\mathbf{E}\nu_{i,t}\nu_{j,t+s}}{b^2} = \frac{\sigma_\alpha^2}{b^2}\left[\frac{\rho^s}{1-\rho^2} + \left(\iota\left(i,j\right) + \frac{\rho^{s+2}}{n}\frac{1}{1-\rho^2}\right)\lambda\right]$$
$$\text{with } \lambda \equiv \frac{\sigma_\mu^2}{\sigma_\alpha^2} \text{ and } \iota\left(i,j\right) = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j \end{cases} \tag{18}$$

With this formula we can use Generalized Least Squares and easily obtained panel data to estimate the model parameters, and then construct point estimates of the two objects of primary interest, k and r. The scaling factor b drops out in the ratios k and r. Table 2 summarizes the results.

$\frac{\widehat{var(b_{0i})}}{b^2}$	$\frac{\sigma_\alpha^2}{b^2}$	$\lambda \equiv \frac{\sigma_\mu^2}{\sigma_\alpha^2}$	k	r

Table 2 Estimates of model parameters and the implied point estimates of k (the correction to the ranking criterion when international trade in permits is prohibited) and r (the measure of the relative importance of unobserved to observed heterogeneity)

7 The non-cooperative choice of policies

Mideksa (2019) considers the equilibrium choice between a tax and a quota when self-interested countries rather than a global planner choose the policy instrument and level. I use his model to illustrate the different policy implications of going from a representative-agent to an n-agent model by either fragmenting the economy or by merely adding agents. The noncooperative setting also provides a rationale for the suboptimality of policy targets. The resulting increase in the social cost of carbon creates another reason for the global planner to prefer quotas.

Mideksa models the many-country setting by adding countries to the one-country model. Aggregate emissions equal $E = e_i + E_{-i}$ when country i emits e_i and the other countries emit E_{-i}. With Mideksa's formulation and my notation, country i's expected payoff is $W_i = B_0 e_i - \frac{B}{2}\left(e_i\right)^2 - \left(D_0 E + \frac{D}{2}E^2\right)$; the global planner's expected payoff is $W = \sum_i W_i$. Country i's slopes of

[15] Appendix C.4 describes an even simpler estimation strategy, and explains why I do not use it here. The simpler strategy might be especially useful with large samples.

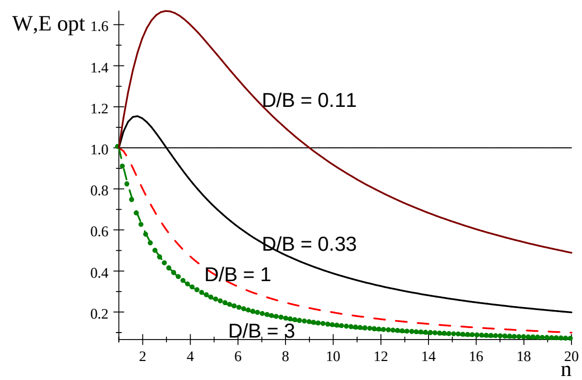

Figure 3: By choice of units I set $D_0 = 0$ and $B_0 = 2$ and I restrict $B + D = B_0$, so that the optimal level of emissions and the maximum expected payoff in the FE equals 1, shown as the horizontal line. The three other graphs show the optimal level of emissions and the maximum expected payoff in the model that increases n by adding agents. These relations depend on n and on $\frac{D}{B}$.

marginal benefit and marginal damages equal B and D respectively. When countries' shocks are uncorrelated (i.e. $\sigma_\alpha^2 = 0$), e_i and E_{-i} are uncorrelated under taxes. Here, i's dominant strategy is to use taxes instead of a quota if and only if $\frac{D}{B} < 1$ – the familiar Weitzman criterion. The global planner takes into account the effect of i's emissions on all n countries, so the slopes for this planner are B and nD. The planner wants to use taxes instead of cap and trade if and only if $\frac{nD}{B} < 1$. Self-interested countries find taxes more attractive than does the planner, because individual countries ignore the risk externality created by the stochastic damages arising from taxes.

To allow for changes in strategic incentives without simultaneously changing the technology, I fragment the damage component of welfare, as I did above for the benefit component. If each country faces damages $d_0 E + \frac{d}{2} E^2$, then aggregate damages equal $n\left(d_0 E + \frac{d}{2} E^2\right)$. Aggregate damages are equal in the RAE and the FE for all emissions levels if and only if $d_0 = \frac{D_0}{n}$; and $d = \frac{D}{n}$. In the FE, by construction, the aggregate expected welfare level for a given level of emissions is independent of n, as is the optimal level of emissions and the maximum expected payoff.

In the interest of exposition, I restrict parameters to satisfy $B_0 = B + D$. This restriction implies that both the optimal level of emissions and the maximum level of welfare equal 1 in the FE, as shown by the horizontal line at 1 in Figure 3. By Definition 1 these levels are independent of n.

Figure 3 also graphs, for different values of $\frac{D}{B}$, the optimal level of emissions (which again equals the maximum level of welfare) in the model that increases n by adding agents. For small $\frac{D}{B}$ these graphs are non-monotonic in n, but in every case the graphs asymptote to 0 as $n \to \infty$. If we interpret an increase in n as scaling up the aggregate economy, holding fixed the absorptive capacity of the atmosphere, then for large n it becomes optimal to send aggregate emissions to zero. In contrast, with the fragmentation model, a change in n merely splinters the aggregate economy, without changing its size (or technology).

As noted above, the planner in the FE prefers taxes over cap and trade if and only if $\frac{D}{B} < 1$. It is evident that with uncorrelated shocks, the individual country prefers taxes over cap and trade if and only if $\frac{d}{b} < 1$. Using the definitions of d and b, this inequality is equivalent to $\frac{D}{B} < n^2$.

Table 3 summarizes this information. Both formulations, increasing n by adding countries or by fragmenting the economy, make the global planner, compared to non-cooperative national planners, less willing to use taxes. The difference in ranking criteria is proportional to n in Mideksa's formulation and it is proportional to n^2 in the fragmentation model. The implied misalignment of incentives in choosing the policy instrument is greater under fragmentation, compared to the alternative that simply adds countries.

ways to increase n	a country's criterion	the global planner's criterion	$\frac{\text{country's criterion}}{\text{planner's criterion}}$
add countries	$\frac{D}{B} < 1$	$\frac{D}{B} < \frac{1}{n}$	n
fragment the model	$\frac{D}{B} < n^2$	$\frac{D}{B} < 1$	n^2

Table 3. The criteria for preferring taxes to quotas when regions' cost shocks are uncorrelated

These remarks assume the existence of an international market for emissions permits. A global planner (or an international agreement) likely finds it much easier to set up this market, compared to individual countries in a noncooperative equilibrium. Quotas are less attractive if the individual countries anticipate that the market will not arise. For this reason also, the lack of international cooperation in choosing the policy type tends to make countries less likely to use quotas.

Mideksa also considers the case where the countries' shocks are positively correlated ($\sigma_{\alpha}^{2} > 0$). When one or more countries use taxes in this situation, those countries generate stochastic emissions, which generate stochastic damages. Given the positive correlation of shocks, an individual country's marginal benefits of emissions are then positively correlated with their marginal damages. Stavins (1996) shows that this positive correlation favors the use of quotas. Appendix C.5 notes that, in the fragmentation model: (i) if $b < d$, the unique Nash equilibrium (in the game where countries choose their policy instrument noncooperatively) is for all countries to use quotas; (ii) for $b > d$ in any Nash equilibrium some (but perhaps not all) countries use taxes.

This discussion of the noncooperative choice of policy levels and instruments assumes that damages arise from flow pollutants. As emphasized above, with stock pollutants the policy ranking depends on the level of the abatement target. Suboptimal targets increase the social cost of carbon, a change that favors quotas. Noncooperatively chosen abatement targets are suboptimal from a global perspective. Suppose that the global planner can determine countries' choice of policy instrument, but countries choose their abatement targets non-cooperatively. If this planner can also set up a global market for trade in emissions permits, the suboptimality of countries' abatement targets encourages the planner to prefer cap and trade over taxes, compared to the scenario where the planner chooses both the policy level and the instrument.[16] In this case, a dynamic model may increase the discrepancy between the planner's and the countries' criteria for ranking policy instruments identified above.

However, a comparison of the scenario in which the planner chooses both the policy instrument and level, with the scenario in which countries choose both the instrument and the level noncooperatively, could reverse that conclusion. The weaker equilibrium policies and resulting increase in countries' social cost of carbon in the noncooperative setting might make them prefer quotas in cases where the global planner prefers taxes.

[16] Proposition 2 involves a convex combination between unregulated emissions and the optimal target. Equilibrium abatement targets in the Markov perfect equilibrium to the noncooperative dynamic game are not a convex combination of those two extreme cases. Therefore, the proposition cannot be applied directly. A formal analysis requires solving the dynamic game. This is easily done, but it would greatly extend the paper's length and complexity, adding minimal insight.

8 Conclusion

International climate negotiations have not succeeded in providing the foundation for an international carbon market. Neither have they resulted in abatement targets that would come close to keeping global mean temperatures below the 2^oC threshold. Reducing emissions and achieving efficiency by means of a market for exchanging abatement credits, are distinct problems. Both are harder in the international, compared to the domestic, setting.

I provide the ranking criteria, with and without trade in emissions permits, for sub-optimal tax and quota pairs. These policy pairs are comparable because – just like the optimal tax and quota – they produce the same level of expected emissions. The setting applies to a global stock pollutant like greenhouse gasses. The gains from trade are generically positive, so it is no surprise that the absence of a market for permits favors taxes. When agents face the same tax, their equilibrium marginal costs are equal, so the allocation of a given amount of abatement is efficient. With a quota, this efficiency requires the existence of a market in permits. Weaker regulation, reflected in lower carbon taxes or higher quotas, increases future emissions, raising the social cost of carbon and its slope, favoring taxes.

I use these results, parameter estimates taken from the literature, and new econometric evidence to examine the choice between taxes and quotas for the regulation of greenhouse gasses. My major results use a period of data consistent with the underlying model. I find that [now summarize empirical results].

A noncooperative setting illustrates the quantitative importance of how we construct a many-agent model based on a one-agent representative model. It is important to "fragment" the one-agent model (as in Weitzman 1974, Section 5), not simply add additional agents. With flow pollutants, agents who choose the policy instrument and level noncooperatively are more likely than the global planner to prefer taxes over quotas.

Stock pollutants might attenuate or magnify this tendency. If, for example, countries choose targets non-cooperatively, but the planner can induce countries to select a particular instrument and also create an international market for permits, the larger social cost of carbon arising from the weak noncooperative equilibrium targets favors quotas. However, if we compare the scenario where the planner chooses both the instrument and the target levels, with the scenario where noncooperative regions choose both, the weaker targets in the latter scenario might lead to a quota instead of a tax.

This possibility is larger if an international carbon market is less likely to arise in the non-cooperative setting.

References

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012): "New trade models: same old gains?," *American Economic Review*, 102, 94–130.

Copeland, B., and M. Taylor (2005): "Free trade and global warming: a trade theory view of the Kyoto Protocol," *Journal of Environmental Economics and Management*, 49(2), 205–234.

Costinot, A., and A. Rodriguez-Clare (2018): "The US Gains from Trade: Valuation using the Demand for Foreign Factor Services," *Journal of Economic Perspectives*, 32(2), 3–24.

Edmonds, J., D. Farrister, L. Clarke, S. de Clara, and C. Munnings (2019): "The Economic Potential of Article 6 of the Paris Agreement and Implementation Challenges," International Emissions Trading Association.

Fally, T., and J. Sayre (2019): "Commodity Trade Matters," NBER working paper.

Hassler, J., P. Krusell, and A. Smith (2017): "Environmental Macroeconomics," in *Handbook of Macroeconomics, vol 2B*, ed. by J. Taylor, and H. Uhlig, pp. 139–187. North Holland.

Hoel, M., and L. Karp (2002): "Taxes Versus Quotas for a Stock Pollutant," *Resource and Energy Economics*, 24, 367 – 384.

Karp, L., and C. Traeger (2019): "Prices versus quantities reassessed," CESifo Working paper.

MacDougall, A., N. Swart, and R. Knutti (2017): "The Uncertainty in the Transient Climate Response to Cumulative CO2 Emissions Arising from the Uncertainty in Physical Climate Parameters," *Journal of Climate*, 30, 813–27.

MATTHEWS, H., N. GILLETT, P. STOTT, AND K. ZICKÂĚELD (2009): "The proportionality of global warming to cumulative carbon emissions.," *Nature*, 459, 829–33.

MEHLING, M., G. METCALF, AND R. STAVINS (2018): "Linking Climate Policies to advance global mitigation," *Science*, 359, 997–998.

MIDEKSA, T. (2019): "Prices vs. Quantities with Multiple Countries," Working Paper.

NEWELL, R., AND W. PIZER (2003): "Regulating Stock Externalitiues Under Uncertainty," *Journal of Environmental Economics and anagement*, 45, 416–32.

NORDHAUS, W. (2013): *The Climate Casino: Risk, Uncertainty, and Economics for a Warming World*. Yale University Press.

PIRIS-CABEZAS, P., R. LUBOWSKI, AND G. LESLIE (2018): "Carbon prices under carbon market scenarios consistent with teh Paris Agreement," Environmental Defense Fund.

RAFEY, W. (2019): "Droughts, deluges, and (river) diversions: Valuing market-based water reallocation," MIT Working Paper.

STAVINS, R. (1996): "Correlated uncertainty and policy instrument choice," *Journal of Environmental Economics and Management*, 30(2), 218–232.

STAVINS, R. (2019): "The Future of U.S. Carbon Markets," in *Environmental and Energy Policy and the Economy*, ed. by M. Kotchen, J. Stock, and C. Wolram, pp. xx–xx. NBER.

WEITZMAN, M. (1974): "Prices versus quantities," *Review of Economic Studies*, 41(4), 477–91.

——— (2012): "GHG target as insurance against catastrophic climate damages," *Journal of Public Economic Theory*, 14(2), 221–44.

A Sketches of proofs

Lemma 1 Equality 10 and the first two equalities in 9 follow immediately from Definition 1; the third equality in 9, arising from ex ante heterogeneity, follows from straightforward but tedious algebra.

Proposition 1 For certainty equivalent policies, in view of Definition 1, moving from the RAE to the FE (with or without trade in permits) does not affect expected damages. Therefore, the proof takes into account only the benefits of emissions. I use the market clearing condition to obtain formulae for the equilibrium quota price (under cap and trade) and each agent's emissions as functions of the arbitrary cap and the realization of the shocks. With this information I calculate the firm's expected gains from trade. Aggregating these gains over firms produces society's expected gains from trade.

I then calculate society's expected benefit of emissions under an arbitrary tax in the FE, showing that this payoff equals the expected payoff in the RAE plus the gains from trade. The tax always results in equality of firms' marginal cost, but the quota requires trade to achieve this equality. Therefore, in the FE, realizing these gains requires a market for permits under the quota, but is automatic under the certainty equivalent tax. The proof makes extensive use of the Principle of Certainty Equivalence, eliminating the need for many tedious calculations.

Lemma 2 I begin with an arbitrary linear policy of the form $m_{0t} + m_1 S_t + m_2 \nu_{t-1}$. With this policy, equation 14 gives emissions under the quota and the certainty equivalent tax; the latter contains the stochastic term arising from the representative agent's response to the shock. For both the tax and the quota, the equilibrium value functions are quadratic in (S_t, ν_{t-1}). Using steps that parallel those in the proof of Proposition 1 in Karp and Traeger (2019), I obtain an expression for the difference in payoffs as a function of the parameters of the model and the coefficients m_1 and m_2. This function has the form $\tilde{g}(m_1, m_2; D, B, \rho, \beta, \delta) \frac{\sigma^2}{2(1-\beta)B}$. Karp and Traeger (2019) provide the formulae for the coefficients of the optimal policy, which I denote here as $m_1(0)$ and $m_2(0)$, functions of the model's parameters. Using the definitions $m_1(c) = (1-c) m_1(0) + c \times 0$ and $m_2(c) = (1-c) m_2(0) + c\frac{\rho}{B}$, I substitute $m_1(c)$ and $m_2(c)$ into $\tilde{g}(\cdot)$ to write $g(c; D, B, \rho, \beta, \delta) \equiv \tilde{g}(m_1(c), m_2(c); D, B, \rho, \beta, \delta)$. Consequently, taxes dominate quotas if and only if $g(c; D, B, \rho, \beta, \delta) > 0$.

Proposition 2 With carbon markets or for $\sigma_\mu^2 = 0$ Lemma 2 provides the policy ranking, leading to statement (b). From Proposition 1.i,

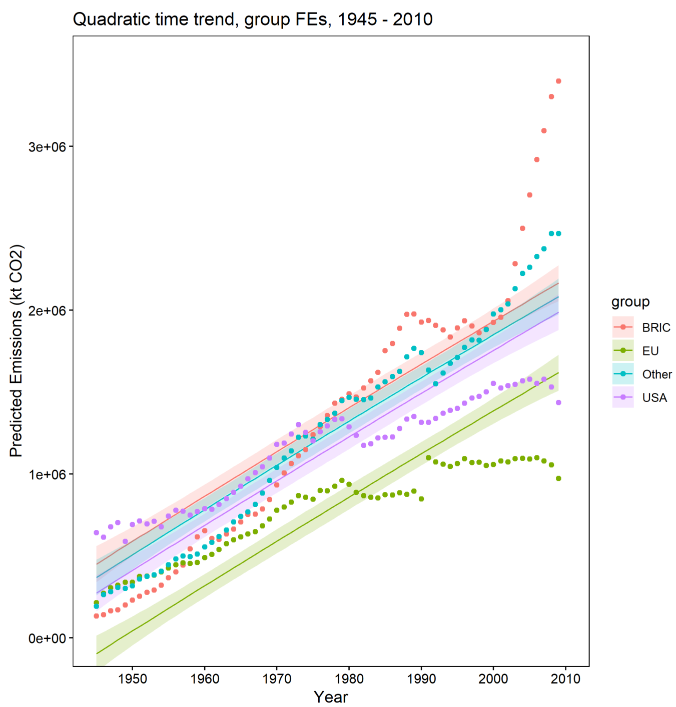

Figure 4: The time series of regions' carbon emissions from 1945 - 2010.

the gains from trade are independent of the quota sequence. Therefore, the expected present discounted stream of expected gains from trade equals $\frac{n-1}{2nB(1-\beta)}\sigma_\mu^2$. The advantage of taxes over quotas-without-trade then equals $\frac{\sigma^2}{2(1-\beta)B}\left[g\left(c; D, B, \rho, \beta, \delta\right) + \frac{n-1}{2n}\frac{\sigma_\mu^2}{\sigma^2}\right]$, resulting in statement (a).

B The regressions

For comparison, Figure 4 shows the graphs of carbon emissions over 1945 – 2010 along with the common time trend. There are two reasons for dropping the final five years of data for the reported results. First, during the last five years emissions in BRIC – primarily in China – grew rapidly. This rapid increase is not consistent with the assumption of a common time trend. Second, the EU Emissions Trading Scheme was phased in during this period, and EU firms anticipated stricter future regulations. Therefore, during this period, EU emissions were somewhat regulated. Given the low price of carbon during this period, the regulation probably had a small effect on carbon emissions.

A more general model replaces equation 17 with

$$e_{it} = \tfrac{1}{b}\left(b_{0i} + \beta x_{it}\right) + \tfrac{1}{b}\nu_{it} = \tfrac{1}{b}\left(b_{0i} + \beta x_{it}\right) + \tfrac{1}{b}\left(\rho\nu_{t-1} + \alpha_t + \mu_{it}\right)$$
$$\text{with } \nu_{it} = \rho\nu_{t-1} + \alpha_t + \mu_{it}.$$

The vector x_{it} can include terms that are common across regions and also region-specific data such as measures of economic activity and of regulation. This model is appropriate to both stock and flow and to regulated and unregulated pollutants. Climate-relevant data involves unregulated emissions that contribute to stock pollutants.

Using equation 16 for the distributional assumptions, and conditions for observational equivalence in equation 9, I write aggregate emissions at t

$$\begin{aligned} e_t &= \textstyle\sum_i e_{it} = \frac{1}{b}\left(\sum_i b_{0i} + \beta \sum_i x_{it}\right) + \frac{1}{b}\sum_i \nu_{it} \\ &= \tfrac{1}{B}\left(B_0 + \beta \bar{x}_t\right) + \tfrac{1}{B}\nu_t \text{ with } \nu_t \equiv \tfrac{\sum_i \nu_{it}}{n} \end{aligned}$$

The aggregate shock is

$$\nu_t \equiv \tfrac{\sum_i \nu_{it}}{n} = \rho\nu_{t-1} + \eta_t$$
$$\text{with } \eta_t \equiv \alpha_t + \theta_t \text{ and } \theta_t \equiv \tfrac{\sum_i \mu_{it}}{n},$$
$$\sigma_\theta^2 = \tfrac{\sigma_\mu^2}{n} \text{ and } \sigma_\eta^2 = \sigma_\alpha^2 + \sigma_\theta^2 = \sigma_\alpha^2 + \tfrac{\sigma_\mu^2}{n}.$$

By repeated substitution

$$\begin{aligned} \nu_t &= \textstyle\sum_{k=0}^{\infty} \rho^k \eta_{t-k} \Rightarrow \\ \nu_{it} &= \left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \eta_{t-k-1} + \alpha_t + \mu_{it}\right) \\ &= \left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\alpha_{t-k-1} + \theta_{t-k-1}\right) + \alpha_t + \mu_{it}\right) \end{aligned}$$

Using the assumption on errors, equation 16, I obtain

$$\mathbf{E}\nu_{it}\nu_{j\tau} = C + D \text{ with}$$
$$C \equiv \mathbf{E}\left[\left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\alpha_{t-k-1}\right) + \alpha_t\right)\left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\alpha_{\tau-k-1}\right) + \alpha_\tau\right)\right]$$
$$D \equiv \mathbf{E}\left[\left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\frac{\sum_i \mu_{it-k-1}}{n}\right) + \mu_{it}\right) \times \left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\frac{\sum_i \mu_{i\tau-k-1}}{n}\right) + \mu_{j\tau}\right)\right].$$

This equation, and straightforward but tedious calculation, produces

$$
\begin{gathered}
\mathbf{E}\nu_{i,t}\nu_{j,t+s} = \frac{\rho^s}{1-\rho^2}\sigma_\alpha^2 + \left(\iota\left(i,j\right) + \frac{\rho^{s+2}}{n}\frac{1}{1-\rho^2}\right)\sigma_\mu^2 \\
\text{with } \iota\left(i,j\right) = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j \end{cases} .
\end{gathered}
\tag{19}
$$

Dividing by b^2 and factoring σ_α^2 produces equation 18, the covariance matrix for the regression equation.

C Referees' Appendix (not for publication)

- Appendix C.1 defines and discusses admissibility.
- Appendix C.2 provides the proofs of Lemma 1 and Proposition 1. Most of the complexity of these proofs arises from the ex ante heterogeneity, which I need only for the empirical application. Readers can simplify the proofs by setting $b_{oi} = B_0$ for all i. I provide an abbreviated proof of Lemma 2 because the full proof parallels the proof of Proposition 1 in Karp and Traeger (2019), as described above in Appendix A. I include the material in the proof below in order to make this paper self-contained. This material provide the expression for the function g in terms of other functions.
- Appendix C.3 provides the formula for the correlation between the permit price and an agent's emissions.
- Appendix C.4 provides a simpler estimation strategy and explains its advantages and disadvantages.
- Appendix C.5 establishes the claim in Section 7 concerning the Nash equilibrium in the policy game when shocks are correlated.

C.1 Admissibility

If the quota is slack, or if the corresponding certainty equivalent tax induces zero emissions, the calculations that underpin the ranking criteria are invalid. Therefore, the linear model provides a reasonable approximation to a more general model only if, with high probability (i) the quota is binding, and (ii)

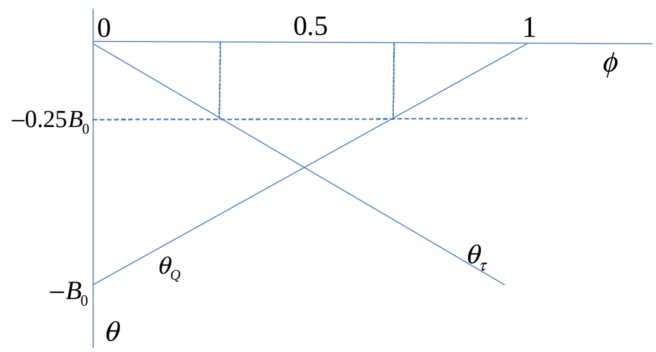

Figure 5: The tax induces positive emissions for $\theta > \theta_\tau$. The quota is binding for $\theta > \theta_Q$. If, for example, the lower bound on θ is $-0.25B_0$, then the quota is binding and the tax induces positive emissions with probabillity 1 ($\alpha = 0$) if and only if $0.25 < \phi < 0.75$.

the tax induces positive emissions. I denote the probability of this event as $1 - \alpha$, with α a modeling choice. For small α, the failure of either condition (i) or (ii) might be acceptable on the grounds that the linear model is viewed as an approximation.

The expected unregulated level of emissions is $\frac{B_0}{B}$. Denote a quota, Q, as a fraction ϕ of unregulated emissions, $Q = \phi\frac{B_0}{B}$. I assume that in expectation the quota is binding, so $\phi < 1$; because the quota is positive, $\phi > 0$. The certainty equivalent tax corresponding to this quota is $\tau = B_0 - B\phi\frac{B_0}{B} = B_0(1-\phi)$. I express the admissibility conditions in terms of ϕ. The quota is slack if and only if $\frac{B_0+\theta}{B} < \phi\frac{B_0}{B}$, i.e. if and only if $\theta < \theta_Q \equiv (\phi - 1) B_0$. The certainty equivalent tax induces zero emissions if and only if $B_0 + \theta < B_0(1-\phi) \Leftrightarrow \theta < \theta_\tau \equiv -\phi B_0$. With these definitions, $\theta_\tau < \theta_Q \Leftrightarrow \phi > 0.5$.

Figure 5 graphs θ_τ and θ_Q as functions of ϕ; the quota is binding if $\theta > \theta_Q$, and the tax induces positive emissions if $\theta > \theta_\tau$. If, for example, the lower bound on θ is $-0.25B_0$, then any quota/tax combination corresponding to $0.25 < \phi < 0.75$ is admissible with $\alpha = 0$. Figure 5 shows these points as the ϕ coordinates of the intersections of the graphs of θ_τ and θ_Q and the horizontal line at $\theta = -0.25B_0$. For a slightly stricter policy ($\phi < 0.25$), condition (ii) fails with positive probability, but condition (i) continues to hold. For a slightly weaker policy ($\phi > 0.75$), condition (i) fails with positive probability, but condition (ii) continues to hold. The Figure illustrates the general point that condition (i) tends to fail (depending on the value of α) for very lax policies, and condition (ii) tends to fail for very strict policies. Increasing α, holding fixed the lower bound on θ, causes the horizontal line to shift up, thereby increasing the range of ϕ that defines the set of admissible policies.

Definition 2 *Denote $\theta_\phi = \max\{\theta_\tau(\phi), \theta_Q(\phi)\}$ and denote $F(\theta)$ as the cumulative distribution function for θ. A tax/quota combination defined by ϕ is admissible if and only if $F(\theta_\phi) \leq \alpha$.*

C.2 The proofs

Proof. (Lemma 1) Part i. I first consider the tax. Facing a tax τ, agent i emits $e_i(\tau) = \frac{b_{0i}-\tau+\theta_i}{b}$. The sum of expected emissions is $\sum_i \frac{b_{0i}-\tau}{b} = \frac{\sum_i b_{0i}-n\tau}{b}$. Setting this expression equal to expected emissions under the tax in the RAE, and equating coefficients of τ, implies the first line of equation 9. The variance of aggregate emissions under a tax in the FE equals

$$\mathbf{E}\left(\frac{\sum_i \theta_i}{b^2}\right)^2 = \mathbf{E}\left(\frac{n\alpha + \sum_i \mu_i}{b^2}\right)^2 = \frac{n^2}{b^2}\left(\sigma_\alpha^2 + \frac{\sigma_\mu^2}{n}\right). \qquad (20)$$

Setting the variance of emissions in the FE under a tax equal to the variance in the RAE, and using the second equality in 9, implies equation 10.

I now consider the quota. Given an aggregate level of emissions, E, information-constrained efficiency requires

$$b_{0i} - be_i = b_{01} - be_1 \forall i.$$

Solving for e_i, summing over i and setting the sum to E, and then solving for e_1, I write i's constrained-efficient expected emissions as

$$e_i^c \equiv \frac{1}{b}\left(b_{0i} - b_{01} - \frac{1}{n}\sum_j (b_{0j} - b_{01})\right) + \frac{E}{n} = \frac{1}{b}(b_{0i} - B_0) + \frac{E}{n}, \qquad (21)$$

and i's expected benefit of emissions as $a + b_{0i}e_i^c - \frac{b}{2}e_2^c$. The next step uses the relation

$$\begin{gathered}\sum_i \left[b_{0i}(b_{0i} - B_0) - \tfrac{1}{2}(b_{0i} - B_0)^2\right] = \\ \sum_i \left[b_{0i}^2 - b_{0i}B_0 - \tfrac{1}{2}(b_{0i}^2 - 2b_{0i}B_0 + B_0^2)\right] = \\ \tfrac{1}{2}\sum_i (b_{0i}^2 - B_0^2) = \tfrac{n}{2}\widehat{var(b_{0i})}\end{gathered} \qquad (22)$$

The last line uses the definition in the second line of equation 9.

Using equation 21, the first equation in 9 and equation 22, the economy-wide expected benefit from the aggregate quota E in the FE without trade equals

$$na + \mathbf{E}\left[\sum_i \left(\left(b_{0i} + \theta_i\right) e_i^c - \frac{b}{2}\left(e_i^c\right)^2\right)\right] =$$

$$na + \mathbf{E}\left[\sum_i \left(\left(b_{0i} + \theta_i\right)\left(\frac{1}{b}\left(b_{0i} - B_0\right) + \frac{E}{n}\right) - \frac{b}{2}\left(\frac{1}{b}\left(b_{0i} - B_0\right) + \frac{E}{n}\right)^2\right)\right] =$$

$$na + \sum_i \left[b_{0i}\left(\frac{1}{b}\left(b_{0i} - B_0\right) + \frac{E}{n}\right) - \frac{b}{2}\left(\frac{1}{b}\left(b_{0i} - B_0\right) + \frac{E}{n}\right)^2\right] =$$

$$na + B_0 E - \frac{B}{2}E^2 + \frac{1}{b}\sum_i \left[b_{0i}\left(b_{0i} - B_0\right) - \frac{1}{2}\left(b_{0i} - B_0\right)^2\right] - \frac{E}{n}\sum_i \left(b_{0i} - B_0\right) =$$

$$na + B_0 E - \frac{B}{2}E^2 + \frac{n}{2b}\widehat{var\left(b_{0i}\right)}.$$

Definition 1.i holds if and only if $a + \frac{1}{2b}\widehat{var\left(b_{0i}\right)} = 0$, implying the second line of equation 9.

Part ii. Absent trade, if each agent receives an equal share of the quota instead of the information-constrained optimal share, the expected benefit of emissions from an aggregate quota E is

$$na + \mathbf{E}\left[\sum_i \left(\left(b_{0i} + \theta_i\right)\frac{E}{n} - \frac{b}{2}\left(\frac{E}{n}\right)^2\right)\right] =$$

$$na + \sum_i \left(\left(b_{0i}\right)\frac{E}{n} - \frac{b}{2}\left(\frac{E}{n}\right)^2\right) = na + \left(\sum_i b_{0i}\right)\frac{E}{n} - \frac{b}{2n}E^2 =$$

$$na + B_0\frac{E}{n} - \frac{B}{2}E^2.$$

Thus, the welfare loss due to equal rather than information-constrained optimal quota shares is

$$-na = \frac{\widehat{var\left(b_{0i}\right)}}{2B}.$$

■

Proof. (Proposition 1) I first establish Parts (i) and (ivb) of Proposition 1. If the planner distributes E emissions permits and allows agents to trade, the market clears at a price, p. Facing price p (which to the agent looks exactly like a tax), agent i emits $e_i(p) = \frac{b_{0,i} - p + \theta_i}{b}$. The definition $\Theta_{-i} \equiv \sum_{j \neq i}\theta_j$

and the market clearing condition, $\sum_i e_i(p) = E$, together with equation 9, imply

$$p = B_0 + \frac{\theta_i + \Theta_{-i}}{n} - BE. \tag{23}$$

The agent's equilibrium level of emissions is

$$e_i(p) = \frac{b_{0i} - p + \theta_i}{b} = e_i^c + \frac{(n-1)}{nb}\left(\theta_i - \frac{\Theta_{-i}}{n-1}\right), \tag{24}$$

where equation 21 gives e_i^c, i's information-constrained efficient level of emissions. The difference between the agent's actual and information-constrained efficient level of emissions is proportional to the difference between the agent's shock and the average shock for other agents.

Using equation 24, i's expected benefit of emissions, given the aggregate quota E, *with trade* in permits, equals

$$\begin{gathered} a + \mathbf{E}\left((b_{0i} + \theta_i) e_i - \tfrac{b}{2} e_i^2\right) = a + b_{0i} e_i^c - \tfrac{b}{2}\left(e_i^c\right)^2 + gft, \text{ with } gft \equiv \\ \mathbf{E}\left[\tfrac{(n-1)}{nb}\theta_i\left(\theta_i - \tfrac{\Theta_{-i}}{n-1}\right) - \tfrac{b}{2}\left(\tfrac{(n-1)}{nb}\right)^2\left(\theta_i^2 - 2\theta_i\tfrac{\Theta_{-i}}{n-1} + \left(\tfrac{\Theta_{-i}}{n-1}\right)^2\right)\right]. \end{gathered} \tag{25}$$

The identity defines the agent's expected gains from trade, gft. *Absent trade*, the agent who receives an information-constrained quota allocation, e_i^c, has the expected benefit of emissions equal to $a + b_{0i} e_i^c - \frac{b}{2}\left(e_i^c\right)^2$.

The second moments needed to calculate gft are:

$$\begin{gathered} \mathbf{E}\left(\theta_i\right)^2 = \sigma_\alpha^2 + \sigma_\mu^2, \ \mathbf{E}\left(\theta_i \Theta_{-i}\right)^2 = (n-1)\,\sigma_\alpha^2, \text{ and} \\ \mathbf{E}\left(\Theta_{-i}\right)^2 = (n-1)^2 \sigma_\alpha^2 + (n-1)\,\sigma_\mu^2. \end{gathered} \tag{26}$$

Using equation 26 and the definition of gft, and simplifying, I obtain:

$$\begin{gathered} gft = \tfrac{(n-1)}{nb}\left(\sigma_\alpha^2 + \sigma_\mu^2\right) - \tfrac{n-1}{nb}\sigma_\alpha^2 - \\ \tfrac{b}{2}\left(\tfrac{(n-1)}{nb}\right)^2\left[\left(\sigma_\alpha^2 + \sigma_\mu^2\right) - 2\sigma_\alpha^2 + \sigma_\alpha^2 + \tfrac{\sigma_\mu^2}{(n-1)}\right] = \tfrac{1}{2}\tfrac{n-1}{nb}\sigma_\mu^2 = \tfrac{1}{2}\tfrac{n-1}{n^2 B}\sigma_\mu^2, \end{gathered}$$

The last equality uses equation 9 to replace the coefficient for the agent in the FE, b, with the coefficient in the RAE, nB. The definition $G(n) \equiv n \times gft$ implies $G(n) = \frac{1}{2}\frac{n-1}{nB}\sigma_\mu^2$, producing equation 11.

Using equations 9, 11 and 25, I write the planner's expected payoff under an aggregate quota E with trade

$$
\begin{aligned}
&na + \sum_i \left(b_{0i}e_i^c - \tfrac{b}{2}\left(e_i^c\right)^2\right) + G = \\
&na + \sum_i \left(b_{0i}\left(\tfrac{1}{b}\left(b_{0i} - B_0\right) + \tfrac{E}{n}\right) - \tfrac{b}{2}\left(\tfrac{1}{b}\left(b_{0i} - B_0\right) + \tfrac{E}{n}\right)^2\right) + G = \\
&na + \sum_i \left(b_{0i}\tfrac{E}{n} - \tfrac{b}{2}\left(\tfrac{E}{n}\right)^2\right) + G \\
&+\tfrac{1}{b}\sum_i \left(b_{0i}\left(b_{0i} - B_0\right)\right) - \tfrac{1}{2b}\sum_i \left(b_{0i} - B_0\right)^2 - \tfrac{E}{nb}\sum_i \left(b_{0i} - B_0\right) + G = \\
&na + B_0E - \tfrac{B}{2}E^2 + \tfrac{1}{b}\sum_i \left(b_{0i}\left(b_{0i} - B_0\right)\right) - \tfrac{1}{2b}\sum_i \left(b_{0i} - B_0\right)^2 + G \\
&= na + B_0E - \tfrac{B}{2}E^2 + \tfrac{1}{2B}\widehat{var\left(b_{0i}\right)} + G = B_0E - \tfrac{B}{2}E^2 + G.
\end{aligned}
\tag{27}
$$

The last equality uses the definition of $\widehat{var\left(b_{0i}\right)}$ in equation 22.

Welfare under a quota-without-trade in the FE equals the last expression in equation 27, minus G. This fact and comparison of equations 1 and 27 imply that the optimal aggregate quota is the same in the RAE and the FE with or without trade, thus establishing Proposition 1 Part (ivb), and also that welfare under the optimal quota is the same in the RAE and in the FE without trade, thus establishing Part (ivc). Given that the optimal quota is the same without or without trade, G indeed represents the change in payoff arising only from trade (not from a change in the policy level), thus establishing Proposition 1 Part (i).

I now establish Proposition 1.iv(d). To this end, I first obtain expressions for the actual and the expected levels of emissions under a tax in the FE, denoted $E(\tau)$ and $\bar{E}(\tau)$, respectively. Agent i's equilibrium response to the tax is

$$e_i(\tau) = \frac{b_{0i} + \theta_i - \tau}{b}.$$

Summing over i and using the first line of equation 9 gives

$$E(\tau) = \frac{\frac{1}{n}\sum_i \left(b_{0i} + \theta_i\right)}{\frac{1}{n}b} - \frac{\tau}{\frac{1}{n}b} = \bar{E}(\tau) + \frac{\sum_i \theta_i}{nB}, \text{ with } \bar{E}(\tau) \equiv \frac{B_0 - \tau}{B}. \tag{28}$$

The aggregate expected benefit of emissions under the tax (plus the constant term) equals

$$
\begin{gathered}
na + \mathbf{E}\left[\sum_i \left(\left(b_{0i} + \theta_i\right) e_i - \frac{b}{2}\left(e_i^2\right)\right)\right] = \\
na + \mathbf{E}\left[\sum_i \left(b_{0i} + \theta_i\right)\left(\frac{b_{0i}+\theta_i-\tau}{b}\right) - \frac{b}{2}\sum_i \left(\frac{b_{0i}+\theta_i-\tau}{b}\right)^2\right].
\end{gathered} \tag{29}
$$

Using equation 28, the first two terms on the right side of equation 29 equal

$$
\begin{gathered}
na + \mathbf{E}\left[\sum_i \left(b_{0i} + \theta_i\right)\left(\frac{b_{0i}+\theta_i-\tau}{b}\right)\right] = na + \frac{n}{b}\left(\frac{\sum_i \left(b_{0i}\left(b_{0i}-\tau\right)\right) + \mathbf{E}\left[\sum_i \theta_i^2\right]}{n}\right) = \\
na + \frac{1}{B}\left[\frac{\left(\sum_i \left(b_{0i}^2\right) - \tau \sum_i \left(b_{0i}\right)\right)}{n} + \mathbf{E}\left(\frac{\sum_i \theta_i^2}{n}\right)\right] = \\
na + \frac{1}{B}\left[\widehat{var\left(b_{0i}\right)} + B_0^2 - \tau B_0 + \sigma_\alpha^2 + \sigma_\mu^2\right].
\end{gathered} \tag{30}
$$

The last line in equation 30 uses the definition of $\widehat{var\left(b_{0i}\right)}$ in equation 9.

The third term on the right side equation 29 is

$$
\begin{gathered}
-\frac{b}{2}\mathbf{E}\left[\sum_i \left(\frac{b_{0i}+\theta_i-\tau}{b}\right)^2\right] = -\frac{1}{2b}\mathbf{E}\left[\sum_i \left(b_{0i} + \theta_i - \tau\right)^2\right] = \\
-\frac{1}{2b}\left[\sum_i \left(b_{0i} - \tau\right)^2\right] - \frac{1}{2b}\mathbf{E}\left[\sum_i \theta_i^2\right] = \\
-\frac{n}{2b}\left[\sum_i \frac{\left(b_{0i}^2 - 2b_{0i}\tau + \tau^2\right)}{n}\right] - \frac{n}{2b}\left(\sigma_\alpha^2 + \sigma_\mu^2\right) = \\
-\frac{1}{2B}\left[\widehat{var\left(b_{0i}\right)} - 2B_0\tau + \tau^2\right] - \frac{1}{2B}\left(\sigma_\alpha^2 + \sigma_\mu^2\right).
\end{gathered} \tag{31}
$$

Using equations 30 and 31, the aggregate expected benefit of emissions under the tax (inclusive of the constant term), is

$$na + \tfrac{1}{B}\left[\widehat{var\left(b_{0i}\right)} + B_0^2 - \tau B_0 + \sigma_\alpha^2 + \sigma_\mu^2\right] -$$

$$\tfrac{1}{2B}\left[\widehat{var\left(b_{0i}\right)} - 2B_0\tau + \tau^2\right] - \tfrac{1}{2B}\left(\sigma_\alpha^2 + \sigma_\mu^2\right) =$$

$$na + \tfrac{1}{2B}\widehat{var\left(b_{0i}\right)} + \tfrac{1}{B}\left[B_0^2 - \tau^2\right] + \tfrac{1}{2B}\left(\sigma_\alpha^2 + \sigma_\mu^2\right) =$$

$$\tfrac{1}{B}\left[B_0^2 - \tau^2\right] + \tfrac{1}{2B}\left(\sigma_\alpha^2 + \sigma_\mu^2\right) =$$

$$\tfrac{1}{B}\left[B_0^2 - \left(B_0 - B\bar{E}\right)^2\right] + \tfrac{1}{2B}\left(\sigma_\alpha^2 + \sigma_\mu^2\right) = B_0\bar{E}_1 - \tfrac{B}{2}\bar{E}^2 + \tfrac{1}{2B}\left(\sigma_\alpha^2 + \sigma_\mu^2\right). \tag{32}$$

Using equation 4, the expected benefit of emissions under a tax in the RAE is

$$B_0\bar{E}_1 - \frac{B}{2}\bar{E}^2 + \frac{\sigma^2}{2B}.$$

By the requirement of observational equivalence, expected damages under a tax are the same in the RAE and the FE. Therefore, the increase in expected payoff under a tax, in moving from the RAE to the FE is

$$\frac{1}{2B}\left(\sigma_\alpha^2 + \sigma_\mu^2 - \sigma^2\right) = \frac{1}{2B}\left(\sigma_\alpha^2 + \sigma_\mu^2 - \left[\sigma_\alpha^2 + \frac{\sigma_\mu^2}{n}\right]\right) = GFT.$$

The first equality uses equation 10 and the second equality uses equation 11. This result establishes Proposition 1.iv(d).

The fact that the payoffs under a tax in the RAE and the FE are the same, apart from a term that is independent of the tax, implies that the optimal tax is the same in the two settings, thus establishing Proposition 1.iv(a).

I now establish Proposition 1.iii. For this purpose, denote $W^{Q,RAE}$, $W^{Q,FE,\text{ no trade}}$ and $W^{Q,FE,\text{ trade}}$, as, respectively, the optimal level of expected welfare under a quota in the RAE, the FE without trade, and the FE with trade. From the previous results we have

$$G(n) = W^{Q,FE,\text{ trade}} - W^{Q,FE,\text{ no trade}} = W^{Q,FE,\text{ trade}} - W^{Q,RAE}. \tag{33}$$

Also, denote $W^{\tau,RAE}$ and $W^{\tau,FE}$ as the optimal expected welfare under the tax in the RAE and the FE, respectively. With these definitions,

$$W^{Q,RAE} > W^{\tau,RAE} \Leftrightarrow W^{Q,RAE} + G > W^{\tau,RAE} + G$$
$$\Leftrightarrow$$
$$W^{Q,FE,\text{ trade}} > W^{\tau,FE},$$

establishing Proposition 1 Part (iii).

Now I establish Proposition 1.iii. With ex post heterogenous agents, absent trade in permits, welfare under the optimal quota is

$$W^{Q,RAE} = W^{Q,FE,\text{no trade}} = B_0 Q^* - \frac{B}{2}Q^{*2} - \left(D_0 Q^* + \frac{D}{2}Q^{*2}\right).$$

Welfare under the optimal tax is

$$W^{\tau,FE} = B_0 \bar{E}^* - \frac{B}{2}\bar{E}^{*2} - \left(D_0 \bar{E}^* + \frac{D}{2}\bar{E}^{*2}\right) + G\left(n\right) + \frac{1}{2B}\left[1 - \frac{D}{B}\right]\sigma^2$$

The fact that $Q^* = \bar{E}^*$ and equation 11 imply that taxes dominate the quota-without-trade if and only if

$$\frac{1}{2B}\left(\frac{\left(n-1\right)\sigma_\mu^2}{n}\right) + \frac{1}{2B}\left[1 - \frac{D}{B}\right]\sigma^2 > 0 \Leftrightarrow \frac{\left(n-1\right)\sigma_\mu^2}{n\sigma^2} + 1 - \frac{D}{B} > 0.$$

■

Proof. (Lemma 2)

The unit of time is arbitrary, so I set it equal to one year. The parameter ϕ equals the number of units of time of each decision period. The parameter α equals the fraction of the current shock that enters firms' current emissions decisions, either in the unregulated scenario or under a tax. These parameters are important in Karp and Traeger (2019) but not here. I set both parameters equal to 1 in the text, but consider the general case in this proof.[17]

Define the state variable $Y_t = \begin{bmatrix} S_t \\ \nu_{t-1} \end{bmatrix}$. Define $J^i\left(m; t, S_t, \nu_{t-1}\right)$ as the value function for $i \in \{\text{tax, quota}\}$ given arbitrary coefficients m_{0t}, m_1 and m_2 of the decision rules in equation 14. These value functions are quadratic in S, ν_{t-1}, $J^i\left(m; t, S_t, \nu_{t-1}\right) = p^i_{0,t} + p'_{1t} Y_t + \frac{1}{2} Y'_t P Y_t$ where P is a 2x2 constant

[17] The parameters f and b in xx are, respectively, B and D in the current paper. Other parameter names are the same.

matrix, $P = -\begin{pmatrix} L & u \\ u & K \end{pmatrix}$. Only p_{0t}^i differs under taxes and quotas (indicated by the index i). In contrast, p_{0t}^i, p_{1t} and P all depend on the decision rule; that is, they depend on m_{0t}, m_1 and m_2, but they are the same for the certainty equivalent tax and quota pair.

Straightforward but lengthy calculations establish that the elements of P equal

$$\begin{aligned} L &\equiv \tfrac{\partial^{2J^i(m;t,S_t,\nu_{t-1})}}{\partial S_t^2} = -\tfrac{B\phi m_1^2 + D\phi}{\beta\delta^2 + 2\beta\delta\phi m_1 + \beta\phi^2 m_1^2 - 1} \\ u &\equiv \tfrac{\partial^{2J^i(m;t,S_t,\nu_{t-1})}}{\partial S_t \partial \nu_{t-1}} = \left(-\tfrac{\beta m_1 m_2 \phi^2 + \beta\delta m_2 \phi}{\beta\delta\rho + \beta\phi\rho m_1 - 1}\right) L + \tfrac{\phi\rho m_1 - B\phi m_1 m_2}{\beta\delta\rho + \beta\phi\rho m_1 - 1}. \end{aligned} \tag{34}$$

Given the arbitrary policy defined by m_{0t}, m_1 and m_2, arguments that parallel the proof of Proposition 1 in Karp and Traeger (2019) (as described in Appendix A) establish that taxes welfare-dominate quotas if and only if

$$g(\cdot) \equiv \alpha - \frac{\beta(L\alpha\phi + 2Bu)}{B} > 0. \tag{35}$$

(Compare to equation 12 in Karp and Traeger (2019).)

I use three expressions taken from Karp and Traeger (2019)

$$\begin{gathered} \varpi = B\left(1 - \beta\delta^2 - \beta\tfrac{D}{B}\phi^2\right),\ \lambda = \tfrac{1}{2\beta\phi}\left(-\varpi + \sqrt{\varpi^2 + 4\beta\phi^2 DB}\right) \\ \mu = \rho\beta\delta\phi\tfrac{\lambda}{B + \beta\phi\lambda - \rho\beta\delta B}. \end{gathered} \tag{36}$$

to write the optimal decision rule. This optimal rule, together with the expression for unregulated emissions, produce the coefficients of "policy c", the convex combination the optimal and unregulated decision rules:

$$m_1 = \beta\lambda\frac{\delta}{B + \beta\lambda\phi}(c-1),\ m_2 = \frac{1}{B}\frac{\rho}{B + \beta\lambda\phi}(B - B\beta\mu + cB\beta\mu + c\beta\lambda\phi). \tag{37}$$

Substituting these expressions into equation 34, and then using these results in inequality 35 gives the ranking criterion under policy c.

To obtain the function $g(\cdot)$, I begin with the model parameters $(\rho, \beta, \delta, D, B)$ and the parameter c that determines the stringency of policy. With these primitives and equations 36 and 37, I obtain the coefficients m_1 and m_2. Substituting these functions in equation 34 I obtain L and u as functions of the model parameters and the policy variable c. Substituting those functions into equation 35 I obtain the function $g(\cdot)$. ■

C.3 Correlation between emissions and the permit price

An agent's emissions are negatively correlated with the emissions price under cap and trade if and only if $n > 2$ and $\sigma_\mu > 0$. This correlation, denoted $\eta_{e,p}(n)$, equals

$$\eta_{e,p}(n) = \chi(n)\frac{\sigma_\mu}{\sigma}, \text{ with } \chi(n) \equiv \frac{-(n-1)^2(n-2)}{n^2(n(n-1))^{0.5}}. \tag{38}$$

For $\sigma_\mu > 0$, a larger n increases the absolute value of the (negative) correlation. For $\sigma_\mu = 0$, agents have the same shock and each agent's non-stochastic share of emissions equals its information-constrained efficient share.

Using equation 10 to write the moments in equation 26 as a function of σ^2 and σ_μ^2, I obtain

$$\begin{gathered}\mathbf{E}(\theta_i)^2 = \sigma_\alpha^2 + \sigma_\mu^2 = \sigma^2 + \tfrac{n-1}{n}\sigma_\mu^2 \\ , \mathbf{E}(\theta_i\Theta_{-i})^2 = (n-1)\sigma_\alpha^2 = (n-1)\left(\sigma^2 - \tfrac{\sigma_\mu^2(n)}{n}\right) \\ \mathbf{E}(\Theta_{-i})^2 = (n-1)^2\sigma_\alpha^2 + (n-1)\sigma_\mu^2 = (n-1)^2\sigma^2 + \tfrac{(n-1)}{n}\sigma_\mu^2.\end{gathered} \tag{39}$$

Equations 23, 24 and 39 imply

$$\begin{gathered}cov(p, e_i) = -\tfrac{1}{n^2}\tfrac{\sigma_\mu^2}{B}(n-1)^2(n-2), \; var(p) = \sigma_\alpha^2 + \tfrac{\sigma_\mu^2}{n} = \sigma^2 \\ var(e_i) = (n-1)\sigma_\mu^2\tfrac{n}{B}.\end{gathered}$$

These results and the definition $\eta \equiv corr(e_i, p)$ imply equation 38. Note that for $n = 2$, $p = B_0 + \frac{\theta_1+\theta_2}{2} - BE$ and $e_1 = \frac{E}{n} + \frac{1}{2b}(\theta_2 - \theta_2)$, so $cov(e_1, p) = \frac{1}{4b}\mathbf{E}((\theta_1+\theta_2)(\theta_1-\theta_2)) = 0$.

The case $n = 2$ is easily understood by means of an example with a two-point distribution and $\sigma_\alpha^2 = 0$. Here, agents' shocks are uncorrelated with each other. The price can take three values: low (when both agents have low shocks), medium (when one agent has a low and the other has a high shock) and high (when both agents have high shocks). An agent's emissions are equal in the low price and the high price states, because the effect of the shock on the agent's demand for emissions exactly offsets the effect of the price. In the medium price states, the agent with a low shock emits less, and the agent with a high shock emits more, than in the other two price states.

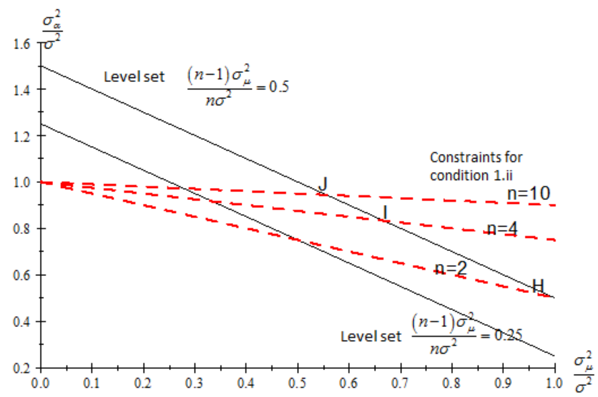

Figure 6: The dashed graphs show the sets of points $\left(\frac{\sigma_\mu^2}{\sigma^2}, \frac{\sigma_\alpha^2}{\sigma^2}\right)$ that satisfy equation 10 (Condition 1.ii) for $n \in \{2, 4, 10\}$. The two other lines show "level sets" for $\frac{(n-1)\sigma_\mu^2}{n\sigma^2}$ equal to 0.5 and 0.25. The points H, I, J identify the combinations of $\left(\frac{\sigma_\mu^2}{\sigma^2}, \frac{\sigma_\alpha^2}{\sigma^2}\right)$ for which $\frac{(n-1)\sigma_\mu^2}{n\sigma^2} = 0.5$ for $n = 2, 4, 10$.

The average level of emissions is the same as in the other two price states. Therefore, the agent's emissions are uncorrelated with the price of emissions.

Figure 6 plots the constraint in equation 10 for $n \in \{2, 4, 10\}$ (the dashed lines). Equation 12 shows that the modification to the slope-based ranking criterion, when there is no trade in permits, equals k. The two solid lines in Figure 6 are (nonstandard) level sets for $k = 0.25$ and $k = 0.5$.[18] For given k, the intersection of a (solid) level set and a (dashed) constraint identify the value of n and the point $\left(\frac{\sigma_\mu^2}{\sigma^2}, \frac{\sigma_\alpha^2}{\sigma^2}\right)$ that produce this value of k. For example, at the points H, I, J, together with the corresponding values $n = 2, 4, 10$, $k = 0.5$. Larger values of the correction, k, cause the level set to move north-east. Thus, a larger value of k is supported by a larger σ_μ^2 and smaller σ_α^2 for given n; equivalently, the same value of k is approximately supported by a larger n, for a given σ_μ^2.[19]

[18] To construct these, hold k fixed and for arbitrary $\frac{\sigma_\mu^2}{\sigma^2}$ solve $k = \frac{n-1}{n}\frac{\sigma_\mu^2}{\sigma^2}$ to obtain $n^* = \frac{\sigma_\mu^2}{\sigma^2}\left(\frac{\sigma_\mu^2}{\sigma^2} - k\right)^{-1}$. Then write the constraint as $\frac{\sigma_\alpha^2}{\sigma^2} = 1 - \frac{\sigma_\mu^2}{n^*\sigma^2} = 1 + k - \frac{\sigma_\mu^2}{\sigma^2}$.

[19] The caveat "approximately" arises because the domain of σ_μ^2 is the positive real line, whereas n is an integer.

C.4 An alternative estimation strategy

An alternative estimation strategy uses the time series of aggregate emissions to estimate ρ and σ^2 and separately uses panel data in which the unit of observation is $e_{it} - \bar{e}_t = e_{it} - \frac{1}{n}\sum_j e_{jt}$ to estimate σ_μ^2. This alternative is simpler because it leads to a closed form expression for the Cholesky decomposition of the relevant covariance matrix, a function only of n. Thus, there is an analytic expression for the inverse of this covariance matrix. In contrast, the covariance matrix defined by equation 19 involves the unknown parameters ρ, σ_α^2, and σ_μ^2, and does not have a closed form Cholesky decomposition. This advantage could be important if nT is large.

However, the alternative presented here separately estimates the sum $\sigma^2 = \sigma_\alpha^2 + \frac{\sigma_\mu}{n}$ and σ_μ^2. With this disjoint estimation, we cannot impose the constraint $\sigma^2 \geq \frac{\sigma_\mu}{n}$, required by $\sigma_\alpha^2 \geq 0$. For the carbon emissions data, the point estimates obtained using the alternative here does not satisfy this inequality. The joint estimation using the procedure described in the text makes it easy to insure that all point estimates are feasible.

Using the first equation in system 16,

$$\frac{1}{n}\sum_i \nu_{it} = \frac{1}{n}\sum_i \left(\nu_t + \mu_{it} - \frac{\sum_j \mu_{jt}}{n}\right) = \nu_t. \tag{40}$$

Using equation 17, average emissions at t equals

$$\bar{e}_t = \frac{1}{n}\sum_i e_{it} = \frac{1}{b}\left(\frac{\sum_i b_{0i}}{n} + h\left(t\right) + \frac{\sum_i \nu_{it}}{n}\right) = \frac{1}{b}\left(B_0 + h\left(t\right) + \nu_t\right), \tag{41}$$

where the second equality uses equations 9 and 40. Lagging equation 41 and multiplying by ρ produces

$$\rho\left(\bar{e}_{t-1} - \frac{1}{b}\left(B_0 + h\left(t-1\right) + \nu_{t-1}\right)\right) = 0.$$

Adding this expression to the right side of equation 41 produces the regression equation

$$\begin{gathered}\bar{e}_t = \rho\bar{e}_{t-1} + \tfrac{1}{b}\left[\left(1-\rho\right)B_0 + h\left(t\right) - \rho h\left(t-1\right)\right] + \omega_t \\ \text{with } \omega_t \equiv \tfrac{\theta_t}{b} \sim iid\left(0, \tfrac{\sigma^2}{b^2}\right).\end{gathered} \tag{42}$$

OLS estimation of this equation provides estimates of ρ, $\frac{\sigma^2}{b^2}$, and of the scaled parameters B_0, c_1 and c_2. An alternative uses GLS to estimate equation 41.

Andy used GLS to estimate

$$E_t = \frac{B_0+h(t)+\nu_t}{B} \text{ with}$$
$$\nu_t = \rho\nu_{t-1} + \theta_t \text{ with } \theta_t \sim iid\left(0,\sigma^2\right).$$

GLS produces more efficient point estimates.

The expressions for $\bar{e}_t$ and e_{it} imply[20]

$$e_{it} - \bar{e}_t = \frac{(b_{0i} - B_0)}{b} + \epsilon_{it} \text{ with } \epsilon_{it} \equiv \frac{1}{b}\left(\mu_{it} - \frac{\sum_i \mu_{it}}{n}\right),\ i = 1, 2...n \qquad (43)$$

Estimating the parameters of equation 43 requires dropping one region (to avoid a singular covariance matrix) and also taking into account the within-period regional correlation arising from the structure of ϵ_{it}. The formulae for the GLS estimator uses the following notation:

- Region i's aggregate emissions over the sample period is $e_i^{AG} = \sum_{t=1}^{T} e_{i,t}$ and the economy-wide aggregate emissions over this period is $n\sum_{t=1}^{T}\bar{e}_t = e^{AG}$;
- I_T is the T dimensional identity matrix and $\otimes$ is the Kronecker product;
- $\mathbf{y}_t = (e_{1,t} - \bar{e}_t, e_{2,t} - \bar{e}_t, ...e_{n-1,t} - \bar{e}_t)'$ is the vector of deviations of region i's emissions from the average emissions in period t and $\mathbf{y} = (\mathbf{y}_1', \mathbf{y}_2', ...\mathbf{y}_T')'$ is the stacked $(n-1)T$ column vector of observations;
- from equation 43, $\mathbf{E}\left(\epsilon_{1t}, \epsilon_{2t}, ...\epsilon_{n-1,t}\right)'\left(\epsilon_{1t}, \epsilon_{2t}, ...\epsilon_{n-1,t}\right) = \frac{\sigma_\mu^2}{nb^2}\Omega$, with the $n-1$ by $n-1$ matrix

$$\Omega = \begin{bmatrix} n-1 & -1 & ... & -1 \\ -1 & n-1 & ... & -1 \\ \vdots & \ddots & \ddots & \ddots \\ -1 & ... & -1 & n-1 \end{bmatrix}.$$

[20]The assumption $\mathbf{E}\left(\alpha\mu_i\right) = 0$ implies that $\sigma^2 \geq \frac{\sigma_\mu^2(n)}{n} \Rightarrow \frac{\sigma^2}{b^2} \geq \frac{\sigma_\mu^2(n)}{nb^2}$

The GLS estimator of $\frac{b_{0i}-B_0}{b}$ is

$$\frac{\widehat{b_{0i}-B_0}}{b}=\frac{e_i^{AG}-\frac{e^{AG}}{n}}{T} \tag{44}$$

and the estimate of the variance is

$$\frac{\widehat{\sigma_\mu^2}}{nb^2}=\frac{\mathbf{y}'\left[\left(I_T-\frac{1}{T}\mathbf{s}\mathbf{s}'\right)\otimes\Omega^{-1}\right]\mathbf{y}}{(n-1)\,T-n} \tag{45}$$

I now provide the details. Estimation of the parameters of equation 43 requires dropping one region to avoid a singular covariance matrix, and then taking into account the within-period correlation of the errors, $\epsilon_{it}=\frac{1}{b}\left(\mu_{i,t}-\frac{\sum_j\mu_{jt}}{n}\right)$. Denote $\tilde{\boldsymbol{\epsilon}}_t=(\epsilon_{1,t},\epsilon_{2,t}...\epsilon_{n,t})'$, the vector of errors at time t, and define the $n\times n$ matrix

$$\tilde{\Omega}=\begin{bmatrix} n-1 & -1 & ... & -1 \\ -1 & n-1 & ... & -1 \\ \vdots & \vdots & \vdots & \vdots \\ -1 & ... & -1 & n-1 \end{bmatrix}.$$

This matrix has $n-1$ on the diagonal and -1 elsewhere. Using the definition of ϵ_{it}, $\mathbf{E}\tilde{\boldsymbol{\epsilon}}_t\tilde{\boldsymbol{\epsilon}}_t'=\frac{\sigma_\mu^2}{nb^2}\tilde{\Omega}$. Because the errors sum to zero, $\tilde{\Omega}$ is singular. Dropping the n'th equation produces $\frac{\sigma_\mu^2}{nb^2}\Omega$, the covariance matrix for $\boldsymbol{\epsilon}_t=(\epsilon_{1,t},\epsilon_{2,t}...\epsilon_{n-1,t})'$, the matrix obtained by dropping the last row and column of $\tilde{\Omega}$. The tilda on $\tilde{\boldsymbol{\epsilon}}_t$ distinguishes the n-dimensional column vector from $\boldsymbol{\epsilon}_t$, the $n-1$-dimensional vector.

Proof. In addition to the notation introduced in Appendix B, I use

- I_{n-1} is the $n-1$ dimensional identity matrix; $\mathbf{s}$ is the T dimensional column vector consisting entirely of 1's.
- $\mathbf{X}=\mathbf{s}\otimes I_{n-1}$, the $(n-1)\,T\times(n-1)$ matrix of stacked $(n-1)\times(n-1)$ identity matrices;
- $\mathbf{f}=(f_1,f_2,...f_{n-1})'$, the column vector of coefficients, with $f_i=\frac{b_{0i}-B_0}{b}$; $\boldsymbol{\epsilon}=(\boldsymbol{\epsilon}_1',\boldsymbol{\epsilon}_2',...\boldsymbol{\epsilon}_T')'$ is the column vector of errors of the $n-1$ included regions.

With this notation, the regression equation 43 becomes

$$\mathbf{y} = X\mathbf{f} + \boldsymbol{\epsilon} \text{ with } \mathbf{E}\boldsymbol{\epsilon} = \mathbf{0} \text{ and } \mathbf{E}\boldsymbol{\epsilon}\boldsymbol{\epsilon}' = (I_T \otimes \Omega)\,\frac{\sigma_\mu^2}{nb^2}. \tag{46}$$

Denote the Cholesky decomposition of Ω as VV'. The relation

$$I_T \otimes \Omega = I_T \otimes (VV') = (I_T \otimes V)(I_T \otimes V') = (I_T \otimes V)(I_T \otimes V)'$$

implies that the Cholesky decomposition of $I \otimes \Omega$ is $(I \otimes V)(I \otimes V)'$. Premultiply the regression 46 by $(I_T \otimes V)^{-1} = (I_T \otimes V^{-1})$ to obtain

$$\left(I_T \otimes V^{-1}\right)\mathbf{y} = \left(I_T \otimes V^{-1}\right)X\mathbf{f} + \left(I_T \otimes V^{-1}\right)\boldsymbol{\epsilon}. \tag{47}$$

The OLS estimator for the transformed system 47 (equivalent to the GLS estimator for the untransformed system 46) is

$$\hat{\mathbf{f}} = \left(X'\left(I_T \otimes \Omega^{-1}\right)X\right)^{-1} X'\left(I_T \otimes \Omega^{-1}\right)\mathbf{y}.$$

To simplify this expression, use

$$\left(X'\left(I_T \otimes \Omega^{-1}\right)X\right)^{-1} = \left((\mathbf{s}'\otimes I_{n-1})\left((I_T \otimes \Omega^{-1})\right)(\mathbf{s}\otimes I_{n-1})\right)^{-1} =$$
$$\left((\mathbf{s}'\otimes\Omega^{-1})(\mathbf{s}\otimes I_{n-1})\right)^{-1} = \left(\mathbf{s}'\mathbf{s}\otimes\Omega^{-1}\right)^{-1} = \tfrac{\Omega}{T}$$

and

$$X'\left(I_T \otimes \Omega^{-1}\right) = (\mathbf{s}'\otimes I_{n-1})\left(I_T \otimes \Omega^{-1}\right) = \mathbf{s}'\otimes\Omega^{-1}.$$

Thus,

$$\hat{\mathbf{f}} = \frac{\Omega}{T}\left(\mathbf{s}'\otimes\Omega^{-1}\right)\mathbf{y} = \frac{1}{T}\left(\mathbf{s}'\otimes I_{n-1}\right)\mathbf{y} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{y}_t, \tag{48}$$

or

$$f_i = \frac{1}{T}\sum_{t=1}^{T}(e_{i,t} - \bar{e}_t) = \quad i = 1.2..n-1.$$

Using the notation introduced above,

$$f_i = \frac{e_i^{AG} - \frac{e^{AG}}{n}}{T},\ i = 1, 2..n-1.$$

As a consistency check, I confirm that the results do not depend on which region is dropped from the regression, recall that $f_i = \frac{b_{0i}-B_0}{b}$. Using equation 9, $\sum_{j=1}^{n} f_j = 0$, or

$$f_n = -\sum_{i=1}^{n-1} f_i = -\sum_{i=1}^{n-1} \frac{e_i^{AG} - \frac{e^{AG}}{n}}{T} =$$

$$-\frac{\left(\frac{n}{n}e^{AG} - e_n^{AG} - \frac{(n-1)}{n}e^{AG}\right)}{T} = \frac{\left(e_n^{AG} - \frac{e^{AG}}{n}\right)}{T}.$$

Thus, I obtain the point estimates of $\frac{b_{0i}-B_0}{b}$ shown in equation 44.

Now I provide the formula for estimating $\frac{\sigma_\mu^2}{nb^2}$. There are $T(n-1)$ observations and n estimated parameters, leaving $(n-1)T - n$ degrees of freedom. Using equations 47 and 48, the vector of residual is

$$\begin{aligned}
\left(I_T \otimes V^{-1}\right)\hat{\boldsymbol{\epsilon}} &= \left(I_T \otimes V^{-1}\right)\mathbf{y} - \left(I_T \otimes V^{-1}\right)X\hat{\mathbf{f}} \\
&= \left(I_T \otimes V^{-1}\right)\left[I_T \otimes I_{n-1} - X\frac{1}{T}\left(\mathbf{s}' \otimes I_{n-1}\right)\right]\mathbf{y} \\
&= \left(I_T \otimes V^{-1}\right)\left[I_T \otimes I_{n-1} - \frac{1}{T}\left(\mathbf{s} \otimes I_{n-1}\right)\left(\mathbf{s}' \otimes I_{n-1}\right)\right]\mathbf{y} \\
&= \left(I_T \otimes V^{-1}\right)\left[I_T \otimes I_{n-1} - \frac{1}{T}\left(\mathbf{ss}' \otimes I_{n-1}\right)\right]\mathbf{y}
\end{aligned}$$

Therefore, sum of squared residuals, SSR, equals

$$\mathbf{y}'\left(\left[I_T \otimes I_{n-1} - \tfrac{1}{T}\left(\mathbf{ss}' \otimes I_{n-1}\right)\right]\right)\left(I_T \otimes V^{-1'}\right)\left(I_T \otimes V^{-1}\right) \times$$
$$\left(\left[I_T \otimes I_{n-1} - \tfrac{1}{T}\left(\mathbf{ss}' \otimes I_{n-1}\right)\right]\right)\mathbf{y}$$

$$= \mathbf{y}'\left(\left[I_T \otimes I_{n-1} - \tfrac{1}{T}\left(\mathbf{ss}' \otimes I_{n-1}\right)\right]\right)\left(I_T \otimes \Omega^{-1}\right) \times$$
$$\left(\left[I_T \otimes I_{n-1} - \tfrac{1}{T}\left(\mathbf{ss}' \otimes I_{n-1}\right)\right]\right)\mathbf{y}$$

$$\begin{aligned}
&= \mathbf{y}'\left[I_T \otimes \Omega^{-1} - \frac{1}{T}\left(\mathbf{ss}' \otimes \Omega^{-1}\right)\right]\left[I_T \otimes I_{n-1} - \frac{1}{T}\left(\mathbf{ss}' \otimes I_{n-1}\right)\right]\mathbf{y} \\
&= \mathbf{y}'\left[I_T \otimes \Omega^{-1} - \frac{1}{T}\left(\mathbf{ss}' \otimes \Omega^{-1}\right) - \frac{1}{T}\left(\mathbf{ss}' \otimes \Omega^{-1}\right) + \frac{1}{T^2}\left(\mathbf{ss}'\mathbf{ss}' \otimes \Omega^{-1}\right)\right]\mathbf{y} \\
&= \mathbf{y}'\left[I_T \otimes \Omega^{-1} - \frac{1}{T}\left(\mathbf{ss}' \otimes \Omega^{-1}\right) - \frac{1}{T}\left(\mathbf{ss}' \otimes \Omega^{-1}\right) + \frac{1}{T}\left(\mathbf{ss}' \otimes \Omega^{-1}\right)\right]\mathbf{y} \\
&= \mathbf{y}'\left[\left(I_T - \frac{1}{T}\mathbf{ss}'\right) \otimes \Omega^{-1}\right]\mathbf{y}.
\end{aligned}$$

Thus, the estimate of $\frac{\sigma_\mu^2}{nb^2}$ is

$$\frac{\sigma_\mu^2}{nb^2} = \frac{\mathbf{y}'\left[\left(I_T - \frac{1}{T}\mathbf{ss}'\right) \otimes \Omega^{-1}\right]\mathbf{y}}{(n-1)\,T - n} \tag{49}$$

■

C.5 Mideksa' model with correlated shocks

To analyze systemic risk ($\sigma_\alpha^2 > 0$), consider matters from the perspective of country i. Denote the set of countries that use emissions taxes as J. Each of those countries' emissions, e_j, $j \in J$, are positively correlated with their shock, which is positively correlated with i's shock. Therefore, these countries' aggregate emissions, $\sum_{j \in J} e_j$, are positively correlated with i's shock. Under the Nash assumption, country i takes the other countries' policies as given. From the perspective of country i, it is as if there is a damage shock that is positively correlated with their shock. From Stavins (1996) we know that this positive correlation favors the use of quotas.

These remarks imply that with fragmentation and $\sigma_\alpha^2 > 0$: (i) if $b < d$, the unique Nash equilibrium is for all countries to use quotas; (ii) for $b > d$ some (but perhaps not all) countries use taxes. To confirm (i), note that if $n-1$ countries use quotas then the remaining country, i, faces no exogenous damage uncertainty. For this case, with $b < d$, we know that i wants to use a quota. Therefore, all countries using quotas is a Nash equilibrium. To confirm that this is the unique Nash equilibrium, suppose to the contrary that there is a Nash equilibrium in which two or more countries use a tax. Taking as given the other countries' policies, any of these tax-setting countries would increase their welfare by deviating to a quota. Therefore, the unique Nash equilibrium is for all countries to use quotas.

To confirm (ii), we note that if $n-1$ countries use quotas, the remaining country, i, faces no damage uncertainty and does strictly better using a tax. Therefore, any Nash equilibrium must have at least one country using a tax. If the variance of the systemic component of the shock is positive but small, all countries use a tax in a Nash equilibrium; if the variance is sufficiently large, a single country uses a tax.

The importance of an international market in emissions permits, discussed in Section 7, also applies when shocks are positively but imperfectly correlated ($0 < \sigma_\alpha^2 < \sigma^2$). With trade in permits (in the model with a flow pollutant), Weitzman's ranking criterion for the planner applies; the planner prefers to use a tax if and only if $B > D$. The tax is even more attractive to the planner without international trade in emissions permits. As noted above, the positive correlation of shocks tends to make taxes less attractive to individual countries operating non-cooperatively. With perfect correlation of shocks, there is no incentive for trade in permits, but the non-cooperative countries' face greater correlation between their shock and (what appears to them as) a damage shock. That greater correlation makes taxes less attractive for the noncooperative countries. The greater correlation does not, however, affect the appeal of the tax for the global planner, because that planner internalizes the randomness of emissions arising from the tax.