Ranking suboptimal climate policies: carbon markets and policy stringency *

Larry Karp†

October 13, 2022

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 almost certainly suboptimal and also likely to be implemented inefficiently. The lack of an international carbon market would favor taxes over quotas; the cost of this missing market dominates familiar factors that effect the two policies' welfare ranking. The suboptimality of abatement targets favors quotas, but the magnitude of this effect is small.

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 of the theory part of this paper. I also thank: Max Auffhammer for suggesting the data; Ethan Ligon and Philippe Boileau for discussions on ML estimation; Aaron Watts for implementing the estimation; Andy Hutgren and Sayantan Mitra for experimenting with other estimation approaches; and Torben Mideksa for conversations on the noncooperative game. The usual disclaimer applies.

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

1 Introduction

The carbon abatement policies recently discussed under the aegis of the United Nations Framework Convention on Climate Change (UNFCCC) are almost certainly suboptimal, and their proposed implementation inefficient. The suboptimality of policy *levels* favors quantity restrictions over taxes, but the lack of an international carbon market favors taxes. Using a simple model, I calculate these offsetting effects and then estimate their magnitudes. The welfare loss resulting from the absence of carbon markets likely swamps the usual considerations that determine whether taxes dominate quantity restrictions. The suboptimality of proposed abatement levels, in contrast, creates only a modest advantage for quantity restrictions.

Signatories to the 2015 Paris Agreement adopted voluntary national carbon abatement targets. Those pledges, 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, Lubowski, & Leslie, 2018). By this measure, proposed policy targets are insufficiently stringent. Due to their different national abatement costs, countries' independent achievement of these targets would be inefficient. Mehling, Metcalf, and Stavins, 2018 and Edmonds et al., 2019 report estimates that an efficient international allocation of abatement could reduce the cost of achieving current targets by 30%, and possibly as much as 75%. That cost reduction would make it easier for countries to adopt the stricter carbon limits needed meet the 2^oC target.

The 2019 Madrid meetings of the of UNFCCC Conference of Parties (COP25), 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 promote efficiency by enabling 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.

The above estimated gains from an international carbon market compare deterministic model outcomes under different trade regimes. In the typical application of this method, used to evaluate trade liberalizaton, we observe the outcome under one regime, the status quo. In the climate setting, where the proposed abatement policies have not been implemented, we observe neither the outcome with or without trade.

There are two additional complications. First, the estimated gains from trade assume particular no-trade national abatement levels, e.g. those consistent with Paris commitments. If countries with high marginal abatement costs are also given to political posturing, the estimated abatement costs in autarchy will be very high, making the gains to trade also high. Second, the reliance on a deterministic model eliminates an important advantage of trade: the fact that markets aggregate information that is either not verifiable or not public.

I address these two issues using a novel method of estimating the gains to an international carbon market. The starting point is that meaningful climate policy will require international agreement. Any negotiated level or trajectory of abatement will likely involve simple policies, e.g. taxes or quantity limits, that have modest informational demands.[1] Tax and quantity restrictions raise similar distributional questions. How should the tax revenue or the quota rents be allocated across signatory countries to achieve political buy-in? I assume that this question has somehow been resolved.

I suppose that the negotiated aggregate abatement level is $c(100)$ percent less than the socially optimal level. Thus, $c = 0$ corresponds to the optimal target and $c \approx 1$ is a weak policy close to Business as Usual (BAU). As noted above, the gains from trade in carbon permits can be made extremely large simply by assuming that the quota allocation is very inefficient. I therefore assume that the allocation is efficient, conditional on c and on public (and verifiable) information. In this setting, the gains from trade arise entirely from the market's ability to aggregate private and non-verifiable information. My estimate of the gains from trade understates actual gains to the extent that the actual quota allocation deviates from the information-constrained optimal allocation.

This machinery also provides a novel perspective on the efficiency ranking of taxes and quotas when there is asymmetric information between the regulator and firms. In the climate setting, the regulator (in my model) corresponds to the agency that implements the climate agreement, summarized by c and the form of policy (a tax or quota). When the agreement specifies an aggregate emissions ceiling, this regulator uses (only) public information to allocate quota *shares* to equate countries' autarchic expected marginal abatement costs; these shares are "information-constrained efficient". A country would want to claim that it has high marginal abatement costs, to receive a high quota allocation. Unless the claim is verifiable, it cannot be used to assign the quota allocation. With efficient markets and international trade in permits, the equilibrium allocation of abatement (as distinct from the initial allocation of permits) does not depend on whether a fact is verifiable.

When comparing a tax and a quota, I restrict the tax to produce the same expected level of emissions as the quota: the two policies are "certainty equivalents".[2] By requiring that quota shares are information-constrained efficient, I achieve an apples-to-apples comparison of policies, subject to assumption that there is not an international market for trade in permits.

The literature comparing taxes and quotas under asymmetric information between the regulator and firms, beginning with Weitzman, 1974, always assumes that quantity restrictions involve trade in permits. Trade weakly increases welfare under a quantity restriction,

[1]The qualifier "likely" in this sentence recognizes that complicated policies or mechanisms could elicit private and non-verifiable information. This important possibility is beyond the scope of this paper.

[2]This restriction does not bind for the optimal tax and quota.

so the absence of trade trivially favors taxes. My measure of the gains from trade quantify the importance of this trade. The usual factors that determine the welfare ranking of taxes and quotas in the climate context include the relative slopes of marginal abatement costs and marginal damages, the persistence of the stock, the discount factor, and serial correlation of private information. My results suggest that the gains from trade swamp these considerations. A global quantity restriction without a well-functioning international market for permits – even with information-constrained efficient allocation of quota shares – is almost certainly less efficient than a uniform global tax that leads to the same expected level of emissions.

The previous literature comparing taxes and quotas assumes that the optimal tax and the optimal quota lead to "interior" outcomes. That is, for all realizations of private information, emissions are positive under the tax, and the quota is binding. Because the quadratic model underlying this analysis is only an approximation, and because the support of the random variable is (to my knowledge) never specified, it is more accurate to say that the policies lead to interior outcomes with "high probability", the exact meaning of which is at the researcher's discretion. I want to consider the possibility that the policy level is suboptimal ($c > 0$). To take advantage of the tractability provided by the quadratic structure, I need to assume that the policy is sufficiently stringent (c is small enough), that with "high probability" the outcome is interior. I denote such policies as "admissible". The assumption that the optimal policy is admissible therefore implies that the admissible set requires an upper bound on c. The quota corresponding to $c \approx 1$ would be close to expected BAU emissions, and therefore would not be binding with high probability. This policy is not admissible.

The quadratic model implies that for all admissible policies, the magnitude of the gains from trade does not depend on the policy stringency (c). This result, *like all others in the quadratic model*, depends on the functional assumptions. These are not defended on the grounds of realism, but because they are associated with an approximation that produces insights that would otherwise not be available. The invariance with respect to policy stringency, of my measure of the gains from trade, makes the measure more useful.

In Weitzman's one-period flow pollutant model, the policy ranking (as distinct from the gains from trade) is also invariant to policy stringency, within the set of admissible policies. Newell and Pizer, 2003 (footnote 14) note that the policy ranking for a stock pollutant does not depend on policy stringency when the regulator uses open loop policies. There, the regulator chooses the sequence of future policies at time zero.[3] Both the one-period flow pollutant model and the multiperiod stock pollutant model with open loop policies are static with

[3] Whether the regulator subsequently revises the policy trajectory initially chosen, i.e. uses "open loop with revision" or sticks with the original trajectory, is irrelevant here. The important point is that in both the open loop and the open loop with revision equilibria, the regulator chooses the current policy as if future policies will follow the trajectory chosen today.

respect to information. The open loop information structure has the advantage of simplicity.

However, the only slightly more complicated feedback structure, where policy at a point in time is conditioned on information available at that time, is arguably more relevant to climate policy. Policymakers today, who condition their decisions on current information, should expect their successors to do the same.[4] For a long-lasting problem like climate change, where we can expect significant changes in information, it is not practical or desirable for today's policymaker to try to commit people in the future to carry out plans made today. Therefore, I consider only feedback policies.

In the feedback (unlike the open loop) setting, the welfare comparison between taxes and quantity restrictions with or without trade in permits does depend on policy stringency. 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 policy. A laxer policy (higher c) implies a higher and steeper social cost of carbon, favoring the quantity-based policy. However, in a reasonable parameterization, this effect is modest.

In summary, building on Weitzman (1974) I propose a measure of the gains from trade given an information-constrained efficient allocation of permits. This estimate captures only the gains arising from markets' ability to aggregate private information, and therefore represents a lower bound on actual gains. These gains are invariant to policy stringency. This fact means that the gains from trade in a multiperiod model, appropriate for stock pollutants such as greenhouse gasses, equal the discounted stream of future per period gains. With a dynamic information structure (feedback policies), as is appropriate for the climate problem, a laxer policy favors the use of quantity restrictions rather than taxes.

Key parameters can be estimated using data on unregulated (BAU) emissions. These estimates, combined with a few parameters taken from the literature, lead to two policy-relevant conclusions: (i) The welfare cost of the missing international market for trade in carbon permits, which reduces welfare under quantity restrictions but not taxes, likely swamps the usual considerations in ranking these two policies. (ii) Relatively weak policies favor quantity restrictions over taxes, but the magnitude of this effect is fairly small. These two conclusions are important in the climate context, where proposed policies are quite weak, and attempts to establish an international market in carbon (let alone more sophisticated markets for trade in other types of abatement measures) have not succeeded.

[4]The feedback structure does not mean that policies can be changed month-to-month or minute-to-minute. For example, we can take the period of commitment to equal the length of a period in the model, which in the climate setting might be five or ten years, i.e. similar to the length of a political cycle.

The next section reviews Weitzman's one-period representative agent model, and formalizes the meaning of "admissible" policies. I then describe the heterogenous agent model needed to introduce trade in permits; that model is observationally equivalent to Weitzman's representative agent model. The next section provides the formula for the gains from trade in this one-period setting, thus quantifying the importance of trade in the welfare comparison between taxes and quotas. I then embed this model in the multiperiod setting needed to consider a stock pollutant. Here I generalize the familiar dynamic model by considering both optimal and suboptimal (certainty equivalent) policies. The following section estimates two (new) parameters using data on unregulated emissions. Those estimates, together with parameter values taken from the literature, and the results form previous sections, produce the conclusions described above.[5]

Additional literature Hoel and Karp, 2002 was the first paper to extend Weitzman's static ranking of taxes and quantities to the stock pollution (dynamic) setting where regulators use feedback policies. Karp and Traeger, 2019 consider a more general problem, also providing a more comprehensive literature review than I offer here. They focus on the case where asymmetric information arises from persistent technology shocks that diffuse slowly through the economy. The profession, following Weitzman's insights, has emphasized the role of relative slopes of marginal abatement costs and marginal damages in ranking taxes and quotas. Karp and Traeger, 2019 show that the persistence of both technology shocks and pollution stocks creates an endogenous correlation between the intercepts of marginal abatement costs and the social cost of carbon – the dynamic analog of static marginal damages. The relative change in these functions' intercepts is as important as their relative slopes in the policy ranking. They conclude that where the information asymmetry arises from persistent technology, quotas dominate taxes under reasonable parameterizations – contrary to the consensus view in the profession.

This strand of the literature assumes that there exist carbon markets, and that policies are information-constrained optimal. In contrast, my research questions address the lack of carbon markets and suboptimal policies. I rely on Karp and Traeger, 2019 for estimates of several parameters, but I depart from their focus on persistent technology shocks. Newbery, 2018 and Stavins, 2020 provide recent surveys of policies to reduce carbon emissions.

I referred above to the use of calibrated models to estimate the effect of trade liberalization by comparing outcomes under different trade regimes. A recent alternative uses gravity-style models to econometrically estimate the gains from trade. Arkolakis, Costinot, and

[5]The next iteration of this working paper will use the fragmented model to consider policy choice in a non-cooperative setting. That material, building on Mideksa, 2019, will show how different approaches of linking a representative agent to a many-agent models modifies policy conclusions.

Rodriguez-Clare, 2012 estimate U.S. gains from trade at 1.4% of 2000 GDP. Costinot and Rodriguez-Clare, 2018 review the literature, reporting estimates of the gains for trade from a large economy such as the US at 2 - 8% of GDP. Fally and Sayre, 2019 disaggregate 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, and many times as large for small countries.

Given the lack of data of economies that have significantly regulated carbon emissions, there seems no hope of using these kinds of models to estimate the gains from trade in carbon permits. However, my approach is closer to these approaches than to the alternative that compares output of deterministic models, insofar as I rely on (very simple) econometrics to recover key parameters.

2 Preliminaries

I describe Weitzman's (1974) model for ranking the optimal tax and quota and then discuss the restriction needed to use this model to compare sub-optimal policies.

There is asymmetric information: firms have more information when they choose emissions than the regulator has when choosing the policy. 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 $\theta \sim (0, \sigma^2)$.[6] Society incurs the pollution cost (external to the polluter) $D_0 E + \frac{D}{2} E^2$. Society's welfare is quadratic in E and linear in θ:

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

This model, a second order approximation of a more general structure, would have little analytic value if it required taking into account realizations of θ at which the tax drives emissions to zero or the quota is not binding. In studying the optimal tax and quota, the literature ignores these possibilities, presumably on the assumption that they are unlikely to arise. I adopt the same simplifying assumption with regard to optimal policies. However, I also consider suboptimal policies. A quota close to the expected BAU level of emissions

[6] The random variable θ corresponds to a demand shock for emissions or an abatement cost shock. Treating actions as emissions, a public bad, or abatement, a public good, are equivalent. To confirm this claim, 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.

would be slack approximately half the time if θ is symmetrically distributed.[7] I cannot consider arbitrarily weak quotas and also ignore the possibility that the quota is slack.

My compromise continues to ignore realizations of θ at which the quota is slack, but restricts policies so that the probability of such events is acceptably small. This restriction accommodates suboptimal policies, but not laissez faire. The idea of "admissibility" formalizes this restriction. The expected unregulated level of emissions is $\frac{B_0}{B}$ and the optimal quota (which equates marginal damages and the expected marginal benefit of emissions) is $\frac{B_0-D_0}{B+D}$. For $0 \leq c \leq 1$, denote $Q^c = (1-c)\frac{B_0-D_0}{B+D} + c\frac{B_0}{B}$, the convex combination of the optimal quota and expected BAU emissions. A smaller value of c corresponds to a stonger policy (a smaller quota). The quota is slack if and only if $\frac{B_0+\theta}{B} < Q^c$, i.e. if and only if $\theta < \theta_c \equiv BQ^c - B_0$. Denote $F(\theta)$ as the cumulative distribution function for θ, so $F(\theta_c)$ is the probability of a state of nature at which the quota defined by c is slack.

I define a policy as admissible if and only if $F(\theta_c)$ is sufficiently small, as judged by the researcher. Hereafter, I assume that all suboptimal policies under consideration are admissible. With this understanding, Weitzman's policy ranking criterion holds for both the optimal and suboptimal policies.

Denote $\bar{E}(\tau)$ as the expectation, conditional on public information, of aggregate emissions when firms respond optimally to an emissions tax τ. A tax, τ, and a quota, Q, are "certainty equivalents" if and only if the expected emissions under the tax equals the (binding) quota: $\bar{E}(\tau) = Q$. The optimal tax and quota are certainty equivalents in this linear quadratic setting. Facing a tax τ, the firm's optimal choice of emissions is $E(\tau,\theta) = \frac{B_0+\theta-\tau}{B}$, so expected emissions is $\bar{E}(\tau) = \frac{B_0-\tau}{B}$; the certainty equivalent tax is $\tau^{CE}(Q) = B_0 - BQ$.

Society's welfare under a tax τ is $W^\tau(E(\tau,\theta),\theta)$ and welfare under a quota Q is $W^Q(Q,\theta)$. For any certainty equivalent admissible policy pair, $\left(Q^c, \tau^{CE}(Q^c)\right)$, including the optimal tax and quota ($c = 0$), the difference between expected welfare under the tax and quota is

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

(Appendix C.1) Taxes welfare-dominate quotas if and only if $B > D$. Weitzman's ranking criterion applies to any certainty equivalent admissible policy pair.

[7]If the optimal tax is nearly certain to produce positive emissions, so will a lower tax. Quotas, but not taxes, create potential problems with using this quadratic model to study suboptimal policies.

3 Heterogenous agents

As in Weitzman, 1974, I create the many-agent version of the Representative Agent Economy (RAE) using a "fragmented economy" (FE), one that splits the RAE into n agents, but without changing the technology, and thus without changing the feasible set. I adopt:

Definition 1. *Quota shares are "information-constrained efficient" if and only if, absent trade, they equalize agents'* ***expected*** *marginal benefit of emissions (marginal abatement cost), conditional on the publicly observed information.*

Full efficiency requires trade in permits, or a similar information-revealing mechanism.

In both the representative agent economy and the fragmented economy (RAE and FE, respectively), firms' payoffs contain a constant, they are quadratic in own emissions (E and e_i), and linear in private information (θ and θ_i); damage is quadratic in aggregate emissions (E and $\sum_i e_i$). The two lines of equation 3 show the two social welfare functions.

$$\begin{aligned} RAE &: A + (B_0 + \theta)\, E - \tfrac{B}{2}E^2 - \left(D_0 E + \tfrac{D}{2}E^2\right) \\ FE &: \textstyle\sum_i \left[a + (b_{0i} + \theta_i)\, e_i - \tfrac{b}{2}e_i^2\right] - \left(D_0 \left(\sum_i e_i\right) + \tfrac{D}{2}\left(\sum_i e_i\right)^2\right). \end{aligned} \tag{3}$$

A firm's marginal abatement cost intercept in the FE, $b_{0i} + \theta_i$, is the sum of a publicly observed component, b_{0i}, and private information, θ_i. The publicly observed component makes it possible to fit the model to data and to assess the welfare gain arising from markets, over and above the gains achieved by an information-constrained efficient allocation of quotas.

Observational equivalence of the RAE and the FE means that, given the same tax or the same aggregate quota, they produce the same aggregate results:

Definition 2. *For $n \geq 1$, observational equivalence means that: (i) For every quota and its information-constrained efficient allocation, the mean and variance of aggregate payoffs are the same in the FE without trade and in the RAE. (ii) For every tax, the mean and variance of aggregate emissions and of the aggregate payoffs are the same in the RAE and the FE.*[8]

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

I now define measures of ex ante (i.e., observable) and ex post (private) heterogeneity.

[8]Replacing Definition 2.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, alters the definition of a in the second part of equation 7, but it does not change the measure of the gains from trade, or anything else of substance. Note that the variance of emissions is zero when the quota is always binding.

The measure of ex ante heterogeneity is

$$\widehat{var}\left(b_{0i}\right) \equiv \left(\frac{1}{n}\sum_i b_{0i}^2\right) - \left(\frac{1}{n}\sum_i b_{0i}\right)^2, \tag{4}$$

the variance of the publicly observed component of firms' intercept of marginal abatement cost. The measure of ex post heterogeneity equals the variance of the firm-specific component of private information. The unobserved component of firm i's marginal abatement cost intercept 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[9]

$$\alpha \sim \left(0, \sigma_\alpha^2\right),\ \mu_i \sim iid\left(0, \sigma_\mu^2\right),\ \text{and } \mathbf{E}\left(\alpha\mu_i\right) = 0. \tag{5}$$

The variance σ_μ^2 is the measure of ex post (private information) heterogeneity.

Normalizing welfare by setting the constant in the RAE $A = 0$, I have[10]

Lemma 1. *(i) Observational equivalence, Definition 2, holds if and only if*

$$\sum_i b_{0i} = nB_0;\ b = nB;\ a = -\frac{\widehat{var}\left(b_{0i}\right)}{2nB}\ \textit{and} \tag{6}$$

$$\sigma^2 = \sigma_\alpha^2\left(n\right) + \frac{\sigma_\mu^2\left(n\right)}{n}. \tag{7}$$

(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 $-na = \frac{1}{2B}\widehat{var}\left(b_{0i}\right)$.

Lemma 1 (i) is trivial if agents are ex ante homogenous, i.e. where $\widehat{var}\left(b_{0i}\right) = 0$. There, where the information-constrained quota is $\frac{E}{n}$, I obtain the first two parts of equation 6 by equating aggregate payoffs in the RAE and the FE under an arbitrary aggregate quota, E:

$$\mathbf{E}_\theta\left(B_0 + \theta\right)E - \tfrac{B}{2}E^2 - \left(D_0E + \tfrac{D}{2}E^2\right) =$$
$$n\mathbf{E}_{\{\theta_i\}}\left[a + \left(b_0 + \theta_i\right)\tfrac{E}{n} - \tfrac{b}{2}\left(\tfrac{E}{n}\right)^2\right] - \left(D_0E + \tfrac{D}{2}E^2\right).$$

Equating coefficients produces the first two parts of equation 6. If agents are ex ante het-

[9] Due to the quadratic structure, results depend on only the first two moments of the distribution, 0 and σ^2. Higher moments affect only the admissibility requirement, defined in Section 2.

[10] The first two equations in system 6 reproduce Weitzman's, 1974 equations 24 and 25. Weitzman does not include ex ante heterogeneity, and therefore does not consider the constant a. The degree of freedom in equation 7 parallels the indeterminacy beneath Weitzman's equation 29. Equation 7 produces a simple parameterization of this indeterminacy.

erogenous and receive information-constrained optimal quota shares, agents with larger b_{0i} have higher marginal abatement costs and receive larger quotas. 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 benefit in the FE without trade is the same as in the RAE. Equation 7 follows from noting that when 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 $e_i(\tau) = \frac{b_{0i}-\tau+\theta_i}{b}$. Aggregate emissions are $\sum_i e_i = \frac{B_0-\tau+\frac{1}{n}\sum_i \theta_i}{B}$, so the variance of aggregate emissions under the tax is

$$\frac{1}{(nB)^2}\mathbf{E}\left(n\alpha + \sum_i \mu_i\right)^2 = \frac{1}{B^2}\left(\sigma_\alpha^2 + \frac{\sigma_\mu^2}{n}\right).$$

Equating the expression on the right to the variance in the RAE (using Definition 2.ii) produces equation 7. There is a continuum of solutions to equation 7; for any non-negative function $\sigma_\alpha^2(n) \leq \sigma^2$, there is a non-negative function $\sigma_\mu^2(n)$ satisfying equation 7.

4 Policy ranking in the Fragmented Economy

Section 2 notes the restriction to admissible policies: those for which the quota is binding 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 assumption requires that the variance of idiosyncratic shocks cannot be arbitrarily large. The assumption is weak, if the agents are countries or regions.

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 (for the quadratic model) 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^*. These facts simplify the proof of the following result.

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

$$G = \frac{n-1}{2nB}\sigma_\mu^2. \tag{8}$$

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

$$k + 1 - \frac{D}{B} > 0 \text{ with } k \equiv \frac{(n-1)\,\sigma_\mu^2}{n\sigma^2}. \tag{9}$$

(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 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.[11] Part (i) shows that the gains from trade, beginning with an information-constrained optimal allocation of shares, depend on n, B, and σ_μ^2, but not on damages, observed heterogeneity, $\{b_{0i}\}$, or the variance of the systemic shock, σ_α^2. Part (ii) shows how the ranking criterion changes without trade in emissions permits. Absent trade, the critical ratio of slopes $\frac{D}{B}$ below which taxes dominate, increases from 1 to $1 + k$.[12]

Both the gains from trade and the ranking criterion are invariant to the certainty equivalent admissible policy pair. This result makes it possible to compare two non-optimal policies that achieve the same expected result.

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 $\frac{\widehat{var(b_{0i})}}{2nB}$ (This claim is a direct consequence of Lemma 1.ii.) Allowing trade in permits achieves the additional gain G. The relative importance of these two sources of heterogeneity is

$$r \equiv \frac{G}{\frac{\widehat{var(b_{0i})}}{2nB}} = \frac{(n-1)\,\sigma_\mu^2}{\widehat{var\,(b_{0i})}}. \tag{10}$$

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.[13]

[11] 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.

[12] Suppose that we hold σ^2 fixed and change n, i.e. we fragment the economy to different degrees. From equation 7, σ_α^2 and/or σ_μ^2 must vary with n. From equations 8 and 9, 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 7 implies $0 \le \sigma_\mu^2 \le n\sigma^2$, so $0 \le k \le n-1$.

[13] Referees' Appendix C.3 provides one additional result: an agent's emissions are negatively correlated with the emissions price under cap and trade if and only if $n > 2$ and $\sigma_\mu > 0$. For $n = 2$, an agent's emissions are positively correlated with the price. The appendix discusses this result.

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.

I examine policy ranking with and without trade in permits for both optimal and suboptimal policies. 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 $2^{o}C$ 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.

In the standard model the marginal benefit of emissions in period t equals $B_0 + h(t) + \nu_t - BE_t$, where ν_t is a shock and $h(t)$ is an exogenous trend. Including this trend is important when estimating or calibrating the model, but it does not change the ranking criteria – another consequence of the Principle of Certainty Equivalence.[14]

Define the stock variable S_t as the difference between the current pollution stock and the damage-minimizing level (e.g. the preindustrial level in the case of greenhouse gasses). This stock obeys the difference equation $S_{t+1} = \delta S_t + E_t$; $0 < \delta \leq 1$ measures the stock persistence.

To introduce persistence of shocks in the RAE, previous papers define the period-t shock as $\nu_t = \rho\nu_{t-1} + \theta_t$, 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. Previous papers assume that there is within-period trade in emissions permits and, like the earlier literature that studies flow pollutants, they compare welfare under the optimal tax and the optimal cap-and-trade policy.

5.1 Suboptimal policies

Here I modify results from the standard linear quadratic stock pollution model in order to compare welfare under certainty equivalent suboptimal policies. 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

[14] A simple generalization replaces the time-dependent function $h(t)$ with a state and time-dependent function $h(x_t, t)$, where x_t contains *exogenous publicly observed* demand and cost shifters such as GNP and technology. The information set at t contains x_t. If x_t follow a stochastic process, the information set at t additionally contains the variables used to predict $x_{t+\tau}$. This extension improves model calibration, but does not change the policy ranking.

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; c is a shorthand for policy. Actual emissions at $t+\tau$, $\tau \geq 0$ under policy c equal

$$E_{t+\tau}^{c,i} = \begin{cases} E_{t+\tau}^c(t, S_t, \nu_{t-1}; c) & \text{for } i = \text{quota} \\ E_{t+\tau}^c(t, S_t, \nu_{t-1}; c) + \frac{\theta_{t+\tau}}{B} & \text{for } i = \text{tax} \end{cases}. \tag{11}$$

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.

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

$$\begin{gathered} J^{c,i}(c; t, S_t, \nu_{t-1}) \equiv \\ \underset{\{\theta_{t+\tau}\}_{\tau=0}^{\infty}}{\mathbf{E}} \sum_{\tau=0}^{\infty} \beta^\tau \left[(B_{0t+\tau} + \nu_{t+\tau}) E_{t+\tau} - \frac{B}{2} E_{t+\tau}^2 - \left(D_0 S_{t+\tau} + \frac{D}{2} S_{t+\tau}^2 \right) \right]., \end{gathered} \tag{12}$$

The state variables (S_t, ν_{t-1}) obey the difference equations given above, and emissions are given by equation 11. The flow payoff (the function in square brackets on the right side of equation 12) is the same as in the static model, except that damages here depend on the pollutant stock, S, not the flow, E. Denote $\Delta \equiv J^{c,tax}(c; t, S_t, \nu_{t-1}) - J^{c,quota}(c; t, S_t, \nu_{t-1})$, the expected advantage of taxes rather than quotas under policy c in the RAE.

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

Lemma 2 generalizes Proposition 1 of Karp and Traeger, 2019, which considers the case of optimal policies ($c = 0$).[15] The function $g(0; D, B, \rho, \beta, \delta)$, corresponding to optimal policy, is simple enough to produce comparative statics, but for $c > 0$ (suboptimal policy) I require numerical methods (Section 6. Therefore, I relegate the formula for $g(c; D, B, \rho, \beta, \delta)$ to Appendix C.2.

5.2 The fragmented economy with suboptimal policies

By using the FE it is possible to show how the ranking between taxes and quotas depends on both the stringency of policy and the availability of carbon markets. The dynamic analog of the static FE sets agent i's marginal benefit of emissions in period t to $b_{0i} + h(t) - be_{it}$ and

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

imposes the first two constraints in equation 6. The period- aggregate shock in the RAE is ν_t and agent i's shock in the FE is $\nu_{i,t}$. Equation 13 collects the distributional assumptions relating the shocks in the RAE and the FE.

$$\begin{gathered}\text{(i) RAE: } \nu_t = \rho\nu_{t-1} + \theta_t \\ \text{(ii) FE: } \nu_{i,t} = \rho\nu_{t-1} + \alpha_t + \mu_{i,t}, \text{ with} \\ \text{(iii) } \alpha_t \sim iid\left(0, \sigma_\alpha^2\right), \mu_{i,t} \sim iid\left(0, \sigma_\mu^2\right), \mathbf{E}\left(\alpha_t\mu_{i,\tau}\right) = 0\forall t, \tau, \text{ and} \\ \text{(iv) } \theta_t = \alpha_t + \frac{\sum_j \mu_{jt}}{n}, \text{ with } var\left(\theta_t\right) \equiv \sigma^2 = \sigma_\alpha^2 + \frac{\sigma_\mu^2}{n}.\end{gathered} \tag{13}$$

The first line of equation 13 defines the autoregressive structure of the shock in the standard RAE, as in Section 5.1. The second line states that agent i's shock in the FE has the same autoregressive structure, except that the agent's innovation consists of a systemic component α_t and an idiosyncratic component $\mu_{i,t}$. Line (iii) states that the innovations are uncorrelated across regions and time. Line (iv) states that the innovation in the RAE equals the systemic innovation in the FE plus the average of idiosyncratic components. The formula for $var\left(\theta_t\right) \equiv \sigma^2$ follows from the previous assumptions in the equation; this formula reproduces equation 7 for the static model. This model satisfies observational equivalence and it accommodates intertemporal correlation of aggregate shocks.

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, given admissibility. Therefore, the present discounted expected value of the gains from trade equals $\frac{G}{1-\beta}$. This fact makes it easy to include the market for permits in the policy ranking. Recall also (from equation 9) that $k \equiv \frac{(n-1)\sigma_\mu^2}{n\sigma^2}$, a measure of the importance of the idiosyncratic shock relative to the aggregate shock

Proposition 2. *(a) Absent international carbon markets and given policy indexed by c, taxes dominate quotas if and only if* $g\left(c; D, B, \rho, \beta, \delta\right) + k > 0$*, 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* $g\left(c; D, B, \rho, \beta, \delta\right) > 0$*.*

6 Empirical application

I use panel data on unregulated emissions to estimate the observed heterogeneity parameters, b_{0i}, and the distributional parameters $\rho, \sigma_\alpha^2, \sigma_\mu^2$. These estimates, combined with estimates of B and D from Karp and Traeger, 2019 and the formulae in previous sections lead to the policy conclusions described in the Introduction.

A region's marginal abatement cost equals 0 under BAU. With Region i's marginal benefit of emissions equal to $b_{0i} + h(t) + \nu_{it} - be_{it}$ (Section 5.2), its BAU emissions equal

$$e_{it} = \frac{1}{b}\left(b_{0i} + h(t) + \nu_{i,t}\right). \tag{14}$$

The fixed effect b_{0i} accounts for observed regional heterogeneity, and ν_{it} is region i's shock in period t. The key assumption is that the slope parameter b is constant over regions and over time. I use annual data of carbon emissions from 1945 to 2004 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.

I do not use a longer time series because there was a rapid growth in Chinese emissions during 2005-2010, inconsistent with the assumption of a common time trend. In addition, the EU Emissions Trading Scheme began to be phased in during this period, and EU firms possibly anticipated stricter future regulations, possibly causing observed emissions to deviate from BAU emissions. A generalization 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 modify the model for use with data on regulated pollutants.) This an other generalizations (if feasible) might be interesting, but the chief empirical virtue of this model is its simplicity. A common time trend with region-fixed effects provide the most parsimonious way to distinguish between public and private (non-verifiable) information. Increasing the set of public information, e.g. by including GNP or measures of technology, open the door to other types of disputes about how to calculate the information-constrained quota allocation.

The distributional assumptions in equation 13 produce a simple expression for the upper triangular elements of the covariance matrix corresponding to equation 14

$$\begin{gathered}\frac{\mathbf{E}\nu_{i,t}\nu_{j,t+s}}{b^2} = \frac{1}{b^2}\left[\rho^s \frac{1}{1-\rho^2}\sigma_\alpha^2 + \left(\left(1-\kappa(s)\right)\iota(i,j) + \kappa(s)\frac{\rho^s}{n} + \frac{\rho^{2+s}}{n}\frac{1}{1-\rho^2}\right)\sigma_\mu^2\right] \\ \text{with indicator functions } \iota(i,j) = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j \end{cases}, \text{ and } \kappa(s) = \begin{cases} 0 \text{ if } s = 0 \\ 1 \text{ if } s \neq 0. \end{cases}\end{gathered} \tag{15}$$

The data consists of BAU emissions, where marginal abatement costs are zero. Therefore, the slope of the marginal cost of abatement, b, is not identified. Consequently, I can estimate the model parameters (apart from ρ) only up to the scaling parameter b. Fortunately, most of the objects of interest involve ratios, from which b cancels.

Assuming normality and the distributional assumptions in equation 13, I estimate the parameters in two steps. First, using equations 6 and 14, I obtain the regression equation for the average per region emissions

$$\frac{\sum_i (e_{it})}{n} = \frac{1}{b}\left(B_0 + h\left(t\right) + \rho\nu_{t-1} + \theta_t\right). \tag{16}$$

Using Generalized Least Squares and equation 16, I obtain an estimate of ρ. I then use Maximum Likelihood to estimate the parameters $\{b_{0i}\}, \sigma_\alpha^2, \sigma_{mu}^2$ (scaled by the unidentified b) conditional on the GLS estimate of ρ. I vary ρ to check for sensitivity.

The point estimates reported here are provisional (and therefore I do not report standard errors). The ML estimates for σ_α^2 andσ_{mu}^2 (but not for $\{b_{0i}\}$ are sensitive to the starting guess for those parameters. Table 1 therefore reports only the parameter estimates and the two ratios of interest, k and r, corresponding to the highest value of the maximized log likelihood. (Future work will address the sensitivity of estimates to starting values.)

$\frac{var(b_{0i})}{b^2}$	$\frac{\sigma_\alpha^2}{b^2}$	$\frac{\sigma_\mu^2}{b^2}$	k	r
$2.9\left(10^{10}\right)$	1.03	16.22	2.4	$1.7\left(10^{-9}\right)$

Table 1 Point estimates of model parameters. k is the correction to the ranking criterion when there is no trade in permits. r is the measure of the importance of unobserved to observed heterogeneity

I take estimates of the parameters B and D from Karp and Traeger, 2019; the marginal abatement cost slope B is calibrated using model runs from DICE. The baseline value of D is calibrated assuming that flow damages are 0 at pre-industrial temperatures and that damages at 2^oC equal one percent of world output. That paper uses the TCRE model in which temperature change is proportional to cumulative carbon emissions since the pre-industrial period; it sets the baseline TCRE parameter to $1.65\frac{^oC}{TtC}$. With this model, $\delta = 1$. A slighlty more pessimistic (higher damages) calibration results in a larger value of D. Karp and Traeger, 2019 focus on the case of persistent technological shocks, where $\rho = 1$. Here I use the estimate $\rho = 0.87$, corresponding to the case where the shocks are amalgams of all private information.

The results from these experiments are as follows:

1. The value of the function g is small relative to the point estimate of k, for all values of c and for both the baseline and the slightly more pessimistic case with higher damages. This result implies that the lack of an international market for carbon would swamp the usual factors that determine the ranking of taxes and quotas.

2. The function g decreases with c: laxer policies favor quantity restrictions over taxes. This effect is negligible for the baseline calibration. For the slighlty more pessimistic calibration, the effect remains small unless c is close to 1. However, policies corresponding to $c \approx 1$ are close to BAU, and therefore unlikely to be admissible. Thus, over the relevant range, the fact that proposed policies are suboptimal has little effect on the ranking of taxes and quotas.

3. The fact that the point estimate of r in Table 1 is negligible means that, relative the gains that could be achieved by basing quota allocations on observables, the additional gains provided by the market are negligible. Nevertheless, as the first bullet point states, the absence of markets would swamp the usual considerations that determine the ranking of taxes and quantity restrictions.

7 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 much harder problems 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 the absence of a market for permits favors taxes, under the assumption that a tax would be uniform across nations. When agents face the same tax, their marginal costs are equal in equilibrium, so abatement is efficient conditional on the aggregate level of emissions. With a quota, this efficiency requires the existence of a market in permits.

I use parameter estimates taken from the literature and new econometric evidence to examine the choice between taxes and quotas for the regulation of greenhouse gasses. The usual factors that determine the efficiency ranking of taxes and quotas for a stock pollutant include relative slopes of marginal abatement costs and damages, the discount factor, the level of serial correlation, and the persistence of the stock. However, the usual comparison assumes that there is cap and trade, which for an international pollutant requires a global market for carbon permits. The loss in welfare under a global quantity restriction, due to the absence of an international carbon market would swamp the usual types of considerations described above, causing a uniform tax to dominate a quantity policy.

This conclusion holds even though I assume that under the quantity policy, the quota allocation is efficient, conditional of public (verifiable) information. The increase in welfare arising from using this constrained efficient allocation (instead of giving each region an equal allocation) is enormous compared to the additional gains that arise from enabling the market to aggregate private information. But trade in permits is nevertheless important enough that it is the overriding consideration in ranking taxes and quantity restrictions.

Weaker regulation, reflected in lower carbon taxes or higher quotas, increases future emissions, raising the social cost of carbon and its slope. I show that weaker regulation favors quantity restrictions. This result may be important in other contexts, where we want to compare suboptimal taxes and quotas. However, in the climate context the effect of suboptimality, on the policy ranking, is small for reasonable calibrations.

References

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

Costinot, A., & Rodriguez-Clare, A. (2018). The us gains from trade: Valuation using the demand for foreign factor services. *Journal of Economic Perspectives*, *32*(2), 3–24.

Edmonds, J., Farrister, D., Clarke, L., de Clara, S., & Munnings, C. (2019). The economic potential of article 6 of the Paris Agreement and implementation challenges, International Emissions Trading association.

Fally, T., & Sayre, J. (2019). Commodity trade matters.

Hoel, M., & Karp, L. (2002). Taxes versus quotas for a stock pollutant. *Resource and Energy Economics*, *24*, 367–384.

Karp, L., & Traeger, C. (2019). Prices versus quantities reassessed, CESifo working paper.

Mehling, M., Metcalf, G., & Stavins, R. (2018). Linking climate policies to advance global mitigation. *Science*, *359*, 997–998.

Mideksa, T. (2019). How (not) to price carbon, Unpublished working paper.

Newbery, D. (2018). Policies for decarbonizing carbon in a liberalized power sector. *Kiel Institute for the World Economy Economics Discussion Papers No 2018-29*.

Newell, R., & Pizer, W. (2003). Regulating stock externalitiues under uncertainty. *Journal of Environmental Economics and Management*, *45*, 416–32.

Piris-Cabezas, P., Lubowski, R., & Leslie, G. (2018). Carbon prices under carbon market scenarios consistent with the Paris agreement, Environmental Defense Fund.

Stavins, R. N. (2020). The future of u.s. carbon-pricing policy. *Environmental and Energy Policy and the Economy*, *1*, 8–64.

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

Appendices

This version of the working paper suppresses all appendices except for the first, because they are extremely long and difficult to read in their current form.

A Sketches of proofs

Lemma 1 Equality 7 and the first two equalities in 6 follow immediately from Definition 2; the third equality in 6, arising from ex ante heterogeneity, follows from straightforward but tedious algebra.

Proposition 1 For certainty equivalent policies, in view of Definition 2, 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 11 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 karp19, 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 $\tilde{g}\left(m_1, m_2; D, B, \rho, \beta, \delta\right) \frac{\sigma^2}{2(1-\beta)B}$. karp19 provide the formulae for the coefficients of the optimal policy, which I denote here as $m_1\left(0\right)$ and $m_2\left(0\right)$, functions of the model's parameters. Using the definitions $m_1\left(c\right) = \left(1-c\right) m_1\left(0\right) + c \times 0$ and $m_2\left(c\right) = \left(1-c\right) m_2\left(0\right) + c\frac{\rho}{B}$, I substitute $m_1\left(c\right)$ and $m_2\left(c\right)$ into $\tilde{g}\left(\cdot\right)$ to write $g\left(c; D, B, \rho, \beta, \delta\right) \equiv \tilde{g}\left(m_1\left(c\right), m_2\left(c\right); D, B, \rho, \beta, \delta\right)$. Consequently, taxes dominate quotas if and only if $g\left(c; D, B, \rho, \beta, \delta\right) > 0$.

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, 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).

C:/Users/karp.GIA330/Dropbox/CapTradeCriterion/graphics/predicte

7.28in7.0396in

Figure A1: The time series of regions' carbon emissions from 1945 - 2010.

B The regressions

For comparison, Figure A1 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 14 with

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

$$\text{with } \nu_{it} = \rho\nu_{t-1} + \alpha_t + \mu_{it}.$$

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. However, this generalization involves time-varying region-specific terms ($\frac{1}{b}\left(b_{0i} + \beta x_{it}\right)$ instead of $\frac{1}{b}b_{0i}$). The theory was developed with region-specific terms that do not vary over time.

C Referees' Appendix (not for publication)

- Appendix C.1 derives Weitzman's criterion, equation 2.
- 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 karp19, 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 derives equation 15 and discusses maximum liklihood estimation.
- Appendix C.5 discusses two alternative estimation strategies.
- Appendix C.7 establishes the claim in Section **??** concerning the Nash equilibrium in the policy game when shocks are correlated.

C.1 Derivation of equation 2

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,

$$\bar{E}(\tau) = \frac{B_0 - \tau}{B} \text{ and } \frac{\sigma^2}{B^2}. \tag{17}$$

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

$$\begin{aligned} W^Q(Q) &\equiv \mathbf{E}_\theta \left[(B_0 + \theta) Q - \tfrac{B}{2}Q^2 - \left(D_0 Q + \tfrac{D}{2}Q^2 \right) \right] \\ &= \underline{B_0 Q - \tfrac{B}{2}Q^2 - \left(D_0 Q + \tfrac{D}{2}Q^2 \right)}. \end{aligned} \tag{18}$$

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

$$\begin{aligned} W^{\tau}\left(\bar{E}\right) \equiv \mathbf{E}_{\theta}[&(B_0 + \theta)\left(\bar{E} + \tfrac{\theta}{B}\right) - \tfrac{B}{2}\left(\bar{E} + \tfrac{\theta}{B}\right)^2 \\ &- \left(D_0\left(\bar{E} + \tfrac{\theta}{B}\right) + \tfrac{D}{2}\left(\bar{E} + \tfrac{\theta}{B}\right)^2\right)]. \end{aligned} \tag{19}$$

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

$$W^{\tau}\left(\bar{E}\right) = \underline{B_0\bar{E} - \frac{B}{2}\bar{E}^2 - \left(D_0\bar{E} + \frac{D}{2}\bar{E}^2\right)} + \frac{1}{2B}\left[1 - \frac{D}{B}\right]\sigma^2. \tag{20}$$

Inspection of equations 18 and 20 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.

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 2 gives the net benefit of moving from the quota to its certainty equivalent tax.

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 6. The variance of aggregate emissions under a tax in the FE equals

$$\mathbf{E}\left(\frac{\sum_i \theta_i}{b^2}\right)^2 = \mathbf{E}\left(\frac{n\alpha + \sum_i \mu_i}{b^2}\right)^2 = \frac{n^2}{b^2}\left(\sigma_\alpha^2 + \frac{\sigma_\mu^2}{n}\right). \tag{21}$$

Setting the variance of emissions in the FE under a tax equal to the variance in the RAE, and using the second equality in 6, implies equation 7.

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

$$e_i^c \equiv \frac{1}{b}\left(b_{0i} - b_{01} - \frac{1}{n}\sum_j (b_{0j} - b_{01})\right) + \frac{E}{n} = \frac{1}{b}(b_{0i} - B_0) + \frac{E}{n}, \tag{22}$$

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

$$\begin{gathered}\sum_i \left[b_{0i}(b_{0i} - B_0) - \tfrac{1}{2}(b_{0i} - B_0)^2\right] = \\ \sum_i \left[b_{0i}^2 - b_{0i}B_0 - \tfrac{1}{2}\left(b_{0i}^2 - 2b_{0i}B_0 + B_0^2\right)\right] = \\ \tfrac{1}{2}\sum_i \left(b_{0i}^2 - B_0^2\right) = \tfrac{n}{2}\widehat{var(b_{0i})}\end{gathered} \tag{23}$$

The last line uses the definition in the second line of equation 6.

Using equation 22, the first equation in 6 and equation 23, the economy-wide expected benefit from the aggregate quota E in the FE without trade equals

$$\begin{gathered}na + \mathbf{E}\left[\sum_i \left((b_{0i} + \theta_i)e_i^c - \tfrac{b}{2}(e_i^c)^2\right)\right] = \\ na + \mathbf{E}\left[\sum_i \left((b_{0i} + \theta_i)\left(\tfrac{1}{b}(b_{0i} - B_0) + \tfrac{E}{n}\right) - \tfrac{b}{2}\left(\tfrac{1}{b}(b_{0i} - B_0) + \tfrac{E}{n}\right)^2\right)\right] = \\ na + \sum_i \left[b_{0i}\left(\tfrac{1}{b}(b_{0i} - B_0) + \tfrac{E}{n}\right) - \tfrac{b}{2}\left(\tfrac{1}{b}(b_{0i} - B_0) + \tfrac{E}{n}\right)^2\right] = \\ na + B_0E - \tfrac{B}{2}E^2 + \tfrac{1}{b}\sum_i \left[b_{0i}(b_{0i} - B_0) - \tfrac{1}{2}(b_{0i} - B_0)^2\right] - \tfrac{E}{n}\sum_i (b_{0i} - B_0) = \\ na + B_0E - \tfrac{B}{2}E^2 + \tfrac{n}{2b}\widehat{var(b_{0i})}.\end{gathered}$$

Definition 2.i holds if and only if $a + \frac{1}{2b}\widehat{var(b_{0i})} = 0$, implying the second line of equation 6.

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

$$\begin{gathered}na + \mathbf{E}\left[\sum_i \left((b_{0i} + \theta_i)\tfrac{E}{n} - \tfrac{b}{2}\left(\tfrac{E}{n}\right)^2\right)\right] = \\ na + \sum_i \left((b_{0i})\tfrac{E}{n} - \tfrac{b}{2}\left(\tfrac{E}{n}\right)^2\right) = na + \left(\sum_i b_{0i}\right)\tfrac{E}{n} - \tfrac{b}{2n}E^2 = \\ na + B_0\tfrac{E}{n} - \tfrac{B}{2}E^2.\end{gathered}$$

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 $e_i\left(p\right) = \frac{b_{0,i}-p+\theta_i}{b}$. The definition $\Theta_{-i} \equiv \sum_{j\neq i}\theta_j$ and the market clearing condition, $\sum_i e_i\left(p\right) = E$, together with equation 6, imply

$$p = B_0 + \frac{\theta_i + \Theta_{-i}}{n} - BE. \tag{24}$$

The agent's equilibrium level of emissions is

$$e_i\left(p\right) = \frac{b_{0i} - p + \theta_i}{b} = e_i^c + \frac{(n-1)}{nb}\left(\theta_i - \frac{\Theta_{-i}}{n-1}\right), \tag{25}$$

where equation 22 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 25, i's expected benefit of emissions, given the aggregate quota E, *with trade* in permits, equals

$$\begin{gathered} a + \mathbf{E}\left(\left(b_{0i}+\theta_i\right)e_i - \tfrac{b}{2}e_i^2\right) = a + b_{0i}e_i^c - \tfrac{b}{2}\left(e_i^c\right)^2 + gft, \text{ with } gft \equiv \\ \mathbf{E}\left[\tfrac{(n-1)}{nb}\theta_i\left(\theta_i - \tfrac{\Theta_{-i}}{n-1}\right) - \tfrac{b}{2}\left(\tfrac{(n-1)}{nb}\right)^2\left(\theta_i^2 - 2\theta_i\tfrac{\Theta_{-i}}{n-1} + \left(\tfrac{\Theta_{-i}}{n-1}\right)^2\right)\right]. \end{gathered} \tag{26}$$

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:

$$\begin{gathered} \mathbf{E}\left(\theta_i\right)^2 = \sigma_\alpha^2 + \sigma_\mu^2, \; \mathbf{E}\left(\theta_i\Theta_{-i}\right)^2 = \left(n-1\right)\sigma_\alpha^2, \text{ and} \\ \mathbf{E}\left(\Theta_{-i}\right)^2 = \left(n-1\right)^2\sigma_\alpha^2 + \left(n-1\right)\sigma_\mu^2. \end{gathered} \tag{27}$$

Using equation 27 and the definition of gft, and simplifying, I obtain:

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

The last equality uses equation 6 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 8.

Using equations 6, 8 and 26, I write the planner's expected payoff under an aggregate quota E with trade

$$\begin{aligned}
&na + \textstyle\sum_i \left(b_{0i}e_i^c - \frac{b}{2}\left(e_i^c\right)^2\right) + G = \\
&na + \textstyle\sum_i \left(b_{0i}\left(\frac{1}{b}\left(b_{0i} - B_0\right) + \frac{E}{n}\right) - \frac{b}{2}\left(\frac{1}{b}\left(b_{0i} - B_0\right) + \frac{E}{n}\right)^2\right) + G = \\
&na + \textstyle\sum_i \left(b_{0i}\frac{E}{n} - \frac{b}{2}\left(\frac{E}{n}\right)^2\right) + G \\
&+\tfrac{1}{b}\textstyle\sum_i \left(b_{0i}\left(b_{0i} - B_0\right)\right) - \frac{1}{2b}\sum_i \left(b_{0i} - B_0\right)^2 - \frac{E}{nb}\sum_i \left(b_{0i} - B_0\right) + G = \\
&na + B_0 E - \tfrac{B}{2}E^2 + \frac{1}{b}\textstyle\sum_i \left(b_{0i}\left(b_{0i} - B_0\right)\right) - \frac{1}{2b}\sum_i \left(b_{0i} - B_0\right)^2 + G \\
&= na + B_0 E - \tfrac{B}{2}E^2 + \frac{1}{2B}\widehat{var\left(b_{0i}\right)} + G = B_0 E - \frac{B}{2}E^2 + G.
\end{aligned} \tag{28}$$

The last equality uses the definition of $\widehat{var\left(b_{0i}\right)}$ in equation 23.

Welfare under a quota-without-trade in the FE equals the last expression in equation 28, minus G. This fact and comparison of equations 1 and 28 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 6 gives

$$E(\tau) = \frac{\frac{1}{n}\sum_i (b_{0i} + \theta_i)}{\frac{1}{n}b} - \frac{\tau}{\frac{1}{n}b} = \bar{E}(\tau) + \frac{\sum_i \theta_i}{nB}, \text{ with } \bar{E}(\tau) \equiv \frac{B_0 - \tau}{B}. \tag{29}$$

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

$$\begin{gathered} na + \mathbf{E}\left[\sum_i \left((b_{0i} + \theta_i) e_i - \frac{b}{2}\left(e_i^2\right)\right)\right] = \\ na + \mathbf{E}\left[\sum_i (b_{0i} + \theta_i)\left(\frac{b_{0i}+\theta_i-\tau}{b}\right) - \frac{b}{2}\sum_i \left(\frac{b_{0i}+\theta_i-\tau}{b}\right)^2\right]. \end{gathered} \tag{30}$$

Using equation 29, the first two terms on the right side of equation 30 equal

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

The last line in equation 31 uses the definition of $\widehat{var\,(b_{0i})}$ in equation 6.

The third term on the right side equation 30 is

$$\begin{gathered} -\frac{b}{2}\mathbf{E}\left[\sum_i \left(\frac{b_{0i}+\theta_i-\tau}{b}\right)^2\right] = -\frac{1}{2b}\mathbf{E}\left[\sum_i (b_{0i} + \theta_i - \tau)^2\right] = \\ -\frac{1}{2b}\left[\sum_i (b_{0i} - \tau)^2\right] - \frac{1}{2b}\mathbf{E}\left[\sum_i \theta_i^2\right] = \\ -\frac{n}{2b}\left[\sum_i \frac{\left(b_{0i}^2 - 2b_{0i}\tau + \tau^2\right)}{n}\right] - \frac{n}{2b}\left(\sigma_\alpha^2 + \sigma_\mu^2\right) = \\ -\frac{1}{2B}\left[\widehat{var\,(b_{0i})} - 2B_0\tau + \tau^2\right] - \frac{1}{2B}\left(\sigma_\alpha^2 + \sigma_\mu^2\right). \end{gathered} \tag{32}$$

Using equations 31 and 32, the aggregate expected benefit of emissions under the tax (inclusive of the constant term), is

$$\begin{gathered} na+\tfrac{1}{B}\left[\widehat{var\left(b_{0i}\right)}+B_0^2-\tau B_0+\sigma_\alpha^2+\sigma_\mu^2\right]- \\ \tfrac{1}{2B}\left[\widehat{var\left(b_{0i}\right)}-2B_0\tau+\tau^2\right]-\tfrac{1}{2B}\left(\sigma_\alpha^2+\sigma_\mu^2\right)= \\ na+\tfrac{1}{2B}\widehat{var\left(b_{0i}\right)}+\tfrac{1}{B}\left[B_0^2-\tau^2\right]+\tfrac{1}{2B}\left(\sigma_\alpha^2+\sigma_\mu^2\right)= \\ \tfrac{1}{B}\left[B_0^2-\tau^2\right]+\tfrac{1}{2B}\left(\sigma_\alpha^2+\sigma_\mu^2\right)= \\ \tfrac{1}{B}\left[B_0^2-\left(B_0-B\bar{E}\right)^2\right]+\tfrac{1}{2B}\left(\sigma_\alpha^2+\sigma_\mu^2\right)=B_0\bar{E}_1-\tfrac{B}{2}\bar{E}^2+\tfrac{1}{2B}\left(\sigma_\alpha^2+\sigma_\mu^2\right). \end{gathered} \tag{33}$$

Using equation 19, 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 7 and the second equality uses equation 8. 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

$$G\left(n\right)=W^{Q,FE,\text{ trade}}-W^{Q,FE,\text{ no trade}}=W^{Q,FE,\text{ trade}}-W^{Q,RAE}. \tag{34}$$

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

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

The fact that $Q^* = \bar{E}^*$ and equation 8 imply that taxes dominate the quota-without-trade if and only if

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

□

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.[16]

Define the state variable $Y_t = \begin{bmatrix} S_t \\ \nu_{t-1} \end{bmatrix}$. Define $J^i(m; t, S_t, \nu_{t-1})$ 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 11. These value functions are quadratic in S, ν_{t-1}, $J^i(m; t, S_t, \nu_{t-1}) = p^i_{0,t} + p'_{1t} Y_t + \frac{1}{2} Y'_t P Y_t$ where P is a 2x2 constant matrix, $P = -\begin{pmatrix} L & u \\ u & K \end{pmatrix}$. Only p^i_{0t} differs under taxes and quotas

[16] The parameters f and b in Karp and Traeger, 2019 correspond, respectively, to B and D in the current paper. Other parameter names are the same.

(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

$$\begin{aligned} L \equiv \frac{\partial^{2J^i(m;t,S_t,\nu_{t-1})}}{\partial S_t^2} &= -\frac{B\phi m_1^2+D\phi}{\beta\delta^2+2\beta\delta\phi m_1+\beta\phi^2 m_1^2-1} \\ u \equiv \frac{\partial^{2J^i(m;t,S_t,\nu_{t-1})}}{\partial S_t \partial \nu_{t-1}} &= \left(-\frac{\beta m_1 m_2\phi^2+\beta\delta m_2\phi}{\beta\delta\rho+\beta\phi\rho m_1-1}\right)L+\frac{\phi\rho m_1-B\phi m_1 m_2}{\beta\delta\rho+\beta\phi\rho m_1-1}. \end{aligned} \tag{35}$$

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

$$g(\cdot) \equiv \alpha - \frac{\beta(L\alpha\phi+2Bu)}{B} > 0. \tag{36}$$

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

I use three expressions taken from Karp and Traeger, 2019

$$\begin{aligned} \varpi = B\left(1-\beta\delta^2-\beta\tfrac{D}{B}\phi^2\right),\ \lambda = \tfrac{1}{2\beta\phi}\left(-\varpi+\sqrt{\varpi^2+4\beta\phi^2 DB}\right) \\ \mu = \rho\beta\delta\phi\frac{\lambda}{B+\beta\phi\lambda-\rho\beta\delta B}. \end{aligned} \tag{37}$$

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:

$$m_1 = \beta\lambda\frac{\delta}{B+\beta\lambda\phi}(c-1),\ m_2 = \frac{1}{B}\frac{\rho}{B+\beta\lambda\phi}(B-B\beta\mu+cB\beta\mu+c\beta\lambda\phi). \tag{38}$$

Substituting these expressions into equation 35, and then using these results in inequality 36 gives the ranking criterion under policy c.

To obtain the function $g(\cdot)$, I begin with the model parameters (ρ,β,δ,D,B) and the parameter c that determines the stringency of policy. With these primitives and equations 37 and 38, I obtain the coefficients m_1 and m_2. Substituting these functions in equation 35 I obtain L and u as functions of the model parameters and the policy variable c. Substituting those functions into equation 36 I obtain the function $g(\cdot)$. $\square$

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

$$\eta_{e,p}(n) = \chi(n)\frac{\sigma_\mu}{\sigma}, \text{ with } \chi(n) \equiv \frac{-(n-1)^2(n-2)}{n^2(n(n-1))^{0.5}}. \tag{39}$$

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 7 to write the moments in equation 27 as a function of σ^2 and σ_μ^2, I obtain

$$\begin{gathered}\mathbf{E}(\theta_i)^2 = \sigma_\alpha^2 + \sigma_\mu^2 = \sigma^2 + \tfrac{n-1}{n}\sigma_\mu^2 \\ , \mathbf{E}(\theta_i\Theta_{-i})^2 = (n-1)\sigma_\alpha^2 = (n-1)\left(\sigma^2 - \tfrac{\sigma_\mu^2(n)}{n}\right) \\ \mathbf{E}(\Theta_{-i})^2 = (n-1)^2\sigma_\alpha^2 + (n-1)\sigma_\mu^2 = (n-1)^2\sigma^2 + \tfrac{(n-1)}{n}\sigma_\mu^2.\end{gathered} \tag{40}$$

Equations 24, 25 and 40 imply

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

$$var(e_i) = (n-1)\sigma_\mu^2\tfrac{n}{B}.$$

These results and the definition $\eta \equiv corr(e_i, p)$ imply equation 39. 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. 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 A2 plots the constraint in equation 7 for $n \in \{2, 4, 10\}$ (the dashed lines). Equation 9 shows that the modification to the slope-based ranking criterion, when there is no trade in

C:/Users/karp.GIA330/Dropbox/CapTradeCriterion/graphics/PYNMCVO0

5.4864in8.4726in

Figure A2: 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 7 (Condition 2.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 $n = 2, 4, 10$.

permits, equals k. The two solid lines in Figure A2 are (nonstandard) level sets for $k = 0.25$ and $k = 0.5$.[17] 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.[18]

C.4 Maximum likelihood estimation

This appendix: (i) derives equation 15, (ii) describes the structure of the covariance matrix, and then (iii) discusses the maximum liklihood estimation. To the best of my knowledge there is no canned program that can estimate the model. To write a program that implements maximum liklihood under the assumption of normality, the coder needs to know how the data is organized. That organization determines the structure of the covariance matrix (item ii). To implement the "method of steepest ascent" algorithm for maximizing the liklihood function, the coder needs derivatives that use somewhat arcane rules of matrix differentiation

[17]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 $\frac{\sigma_\alpha^2}{\sigma^2} = 1 - \frac{\sigma_\mu^2}{n^*\sigma^2} = 1 + k - \frac{\sigma_\mu^2}{\sigma^2}$.

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

(item iii).

The aggregate shock in period t is

$$\begin{aligned}\nu_t \equiv \frac{\sum_i \nu_{it}}{n} = \rho\nu_{t-1} + \eta_t \\ \text{with } \eta_t \equiv \alpha_t + \theta_t \text{ and } \theta_t \equiv \frac{\sum_i \mu_{it}}{n}, \\ \sigma_\theta^2 = \frac{\sigma_\mu^2}{n} \text{ and } \sigma_\eta^2 = \sigma_\alpha^2 + \sigma_\theta^2 = \sigma_\alpha^2 + \frac{\sigma_\mu^2}{n}.\end{aligned} \tag{41}$$

I use the indicator functions

$$\iota(i,j) = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j \end{cases}, \text{ and } \kappa(s) = \begin{cases} 0 \text{ if } s = 0 \\ 1 \text{ if } s \neq 0 \end{cases}$$

C.4.1 Deriving equation 15

I want to write the shocks as functions of the α's and the μ's, so that the covariance is a function of the primitives ρ, σ_α^2 and σ_μ^2. By repeated substitution, and assuming assuming $\rho < 1$,

$$\begin{aligned}\nu_t = \textstyle\sum_{k=0}^{\infty} \rho^k \eta_{t-k} \Rightarrow \\ \nu_{it} = \left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \eta_{t-k-1} + \alpha_t + \mu_{it}\right) \\ = \left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\alpha_{t-k-1} + \theta_{t-k-1}\right) + \alpha_t + \mu_{it}\right) \\ = \left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\alpha_{t-k-1} + \frac{\sum_i \mu_{i,t-k-1}}{n}\right) + \alpha_t + \mu_{it}\right)\end{aligned} \tag{42}$$

Using the assumption on errors, equation 41, I obtain

$$\begin{aligned}\mathbf{E}\nu_{it}\nu_{j\tau} = C + D \text{ with} \\ C \equiv \mathbf{E}\left[\left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\alpha_{t-k-1}\right) + \alpha_t\right)\left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\alpha_{\tau-k-1}\right) + \alpha_\tau\right)\right] \\ D \equiv \mathbf{E}\left[\left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\frac{\sum_m \mu_{m,t-k-1}}{n}\right) + \mu_{it}\right) \times \left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\frac{\sum_m \mu_{m,\tau-k-1}}{n}\right) + \mu_{j\tau}\right)\right].\end{aligned} \tag{43}$$

The index m in the sum $\sum_m \mu_{m,t-k-1}$ runs from 1 to n, the number of regions. Equation 43 uses the independence of the α's and μ's, so the expectation of the cross product term in $\mathbf{E}\nu_{it}\nu_{j\tau}$ equals zero. I next take expectations to express the functions C and D in terms of primitives. I use the fact that the covariance matrix is symmetric, so it is sufficient to calculate the covariances only for $\tau \geq t$. Define $s = \tau - t \geq 0$.

For the function C, use

$$\begin{aligned}\left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\alpha_{t-k-1}\right) + \alpha_t\right) = \left(\textstyle\sum_{k=0}^{\infty} \rho^k \alpha_{t-k}\right) \text{ and} \\ \left(\rho \textstyle\sum_{k=0}^{\infty} \rho^k \left(\alpha_{\tau-k-1}\right) + \alpha_\tau\right) = \left(\textstyle\sum_{k=0}^{\infty} \rho^k \alpha_{\tau-k}\right)\end{aligned}$$

to write

$$
\begin{aligned}
C = \mathbf{E}\left[\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{\tau-k}\right)\right] = \\
\mathbf{E}\left[\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t+s-k}\right)\right] = \\
\mathbf{E}\left[\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\textstyle\sum_{k=0}^{s-1}\rho^k\alpha_{t+s-k}+\textstyle\sum_{k=s}^{\infty}\rho^k\alpha_{t+s-k}\right)\right] = \\
\mathbf{E}\left[\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\textstyle\sum_{k=s}^{\infty}\rho^k\alpha_{t+s-k}\right)\right] = \\
\mathbf{E}\left[\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\rho^s\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\right] = \\
\rho^s\sigma_\alpha^2\tfrac{1}{1-\rho^2}.
\end{aligned}
\tag{44}
$$

The third equality in the sequence 44 uses two facts. First, for $s = 0$, $\sum_{k=0}^{s-1}\rho^k\alpha_{t+s-k} = 0$; this is just a convention concerning the meaning of summations. Second, for $s \geq 1$, $\mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\sum_{k=0}^{s-1}\rho^k\alpha_{t+s-k}\right)\right] = 0$ because the two sums do not contain any α's with the same time subscript, and the α's in different periods are uncorrelated. The fourth equality merely changes the definition of the index in the second summation, so that the index runs from 0 instead of from s. The final equality uses the assumption that α_t is iid, so the expectation of all cross-product terms equals zero, resulting in

$$
\begin{aligned}
\mathbf{E}\left[\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\rho^s\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\right] = \\
\rho^s\mathbf{E}\left(\textstyle\sum_{k=0}^{\infty}\rho^{2k}\left(\alpha_{t-k}\right)^2\right) = \rho^s\sigma_\alpha^2\tfrac{1}{1-\rho^2}.
\end{aligned}
$$

For the function D, use

$$
\begin{aligned}
D = \mathbf{E}\left[\left(\mu_{it}+\rho\textstyle\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m\mu_{m,t-k-1}}{n}\right)\right)\left(\mu_{j\tau}+\rho\textstyle\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m\mu_{m,\tau-k-1}}{n}\right)\right)\right] \\
= \mathbf{E}\left[\left(\mu_{it}\right)\left(\mu_{j\tau}\right)\right] + \mathbf{E}\mu_{it}\rho\textstyle\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m\mu_{m,\tau-k-1}}{n}\right) + \\
\mathbf{E}\rho\mu_{j\tau}\textstyle\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m\mu_{m,t-k-1}}{n}\right) + \mathbf{E}\rho\textstyle\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m\mu_{m,t-k-1}}{n}\right)\rho\textstyle\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m\mu_{m,\tau-k-1}}{n}\right).
\end{aligned}
\tag{45}
$$

Denote the four terms after the four expectations operators as D_i with $1 = 1, 2, 3, 4$. I evaluate them in turn. It is immediate that

$$
D_1 = \mathbf{E}\left[\left(\mu_{it}\right)\left(\mu_{j\tau}\right)\right] = \iota\left(i,j\right)\left[1-\kappa\left(s\right)\right]\sigma_\mu^2. \tag{46}
$$

Next, we have

$$D_2 = \mathbf{E}\left[\mu_{it}\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m \mu_{m,\tau-k-1}}{n}\right)\right] = \mathbf{E}\left[\mu_{it}\rho\sum_{k=0}^{\infty}\rho^k\frac{\mu_{i,\tau-k-1}}{n}\right]$$
$$= \mathbf{E}\left[\mu_{it}\rho\sum_{k=0}^{\infty}\rho^k\frac{\mu_{i,t+s-k-1}}{n}\right] = \begin{cases} 0 \text{ if } s=0 \\ \sigma_\mu^2\frac{\rho^s}{n} \text{ if } s\geq 1 \end{cases}. \tag{47}$$

The first equality is a definition; the second uses that fact that $\mathbf{E}\mu_{it}\mu_{j\tau}=0$ for $i\neq j$; and the third uses $\tau=t+s$ from the definition of s. The final equality uses two facts. First, $\mathbf{E}\left[\mu_{it}\rho\sum_{k=0}^{\infty}\rho^k\frac{\mu_{i,t+s-k-1}}{n}\right]=0$ if $s=0$, because in this case the highest time index in the summation is $t-1<t$, so all terms in the sum are uncorrelated with μ_{it}. Second, for $s\geq 1$ the index $k=s-1$ returns μ_{it}, so the expectation for $s\geq 1$ is

$$\mathbf{E}\left[\mu_{it}\rho\sum_{k=0}^{\infty}\rho^k\frac{\mu_{i,t+s-k-1}}{n}\right] = \frac{\rho}{n}\rho^{s-1}\sigma_\mu^2 = \frac{1}{n}\rho^s\sigma_\mu^2.$$

Thus, $D_2=\kappa(s)\frac{1}{n}\rho^s\sigma_\mu^2$.

The third expectation in equation 45 is

$$D_3 = \mathbf{E}\rho\mu_{j\tau}\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m \mu_{m,t-k-1}}{n}\right)$$
$$\mathbf{E}\rho\mu_{j,t+s}\sum_{k=0}^{\infty}\rho^k\left(\frac{\mu_{j,t-k-1}}{n}\right)=0.$$

The first equality is a definition, and the second uses the assumption that the μ's are uncorrelated across regions. The third equality uses the fact that the highest time index in the sum is $t-1<t+s$ because of the assumption that $s\geq 0$. Therefore, $\mu_{j,t+s}$ does not appear in the sum. This step highlights the role of the definition that $\tau=t+s$ with $s\geq 0$. If instead I had written $t=\tau+s$ with $s\geq 0$, the expressions for D_2 and D_3 would have been reversed. Because the covariance matrix is symmetric, it does not matter whether we obtain formulae for the upper or the lower triangular part of the matrix.

For the final expectation, we have

$$D_4 = \mathbf{E}\left[\rho\left(\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m \mu_{m,t-k-1}}{n}\right)\right)\rho\left(\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_m \mu_{m,\tau-k-1}}{n}\right)\right)\right]$$
$$= \frac{\rho^2}{n^2}\mathbf{E}\left[n\left(\sum_{k=0}^{\infty}\rho^k\mu_{m,t-k-1}\right)\left(\sum_{k=0}^{\infty}\rho^k\mu_{m,\tau-k-1}\right)\right] \tag{48}$$
$$= \frac{\rho^2}{n}\mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\mu_{m,t-k-1}\right)\left(\sum_{k=0}^{\infty}\rho^k\mu_{m,t+s-k-1}\right)\right].$$

The first equality is a definition. The second factors out the $\frac{\rho^2}{n^2}$ and uses the fact that the μ's are uncorrelated across regions, so the expectation of all terms involving shocks in different regions vanishes. For each region we are left with the expectation of $\left(\sum_{k=0}^{\infty}\rho^k\mu_{m,t-k-1}\right)\left(\sum_{k=0}^{\infty}\rho^k\mu_{m,\tau-k-1}\right)$, and there are n of these regions. The third equality cancels the n's and uses the definition $\tau = t + s$ with $s \geq 0$. Now use

$$\sum_{k=0}^{\infty} \rho^k \mu_{m,t+s-k-1} = \sum_{k=0}^{s-1} \rho^k \mu_{m,t+s-k-1} + \sum_{k=s}^{\infty} \rho^k \mu_{m,t+s-k-1}$$
$$= \sum_{k=0}^{s-1} \rho^k \mu_{m,t+s-k-1} + \rho^s \sum_{k=0}^{\infty} \rho^k \mu_{m,t-k-1}.$$

Here, the second equality follows from redefining the summation index. Using this result in equation 48 produces

$$\begin{aligned} D_4 &= \tfrac{\rho^2}{n} \mathbf{E} \left[\left(\sum_{k=0}^{\infty} \rho^k \mu_{m,t-k-1} \right) \left(\sum_{k=0}^{s-1} \rho^k \mu_{m,t+s-k-1} + \rho^s \sum_{k=0}^{\infty} \rho^k \mu_{m,t-k-1} \right) \right] \\ &= \tfrac{\rho^2}{n} \mathbf{E} \left[\left(\sum_{k=0}^{\infty} \rho^k \mu_{m,t-k-1} \right) \left(\rho^s \sum_{k=0}^{\infty} \rho^k \mu_{m,t-k-1} \right) \right] \\ &= \tfrac{\rho^{2+s}}{n} \mathbf{E} \left[\left(\sum_{k=0}^{\infty} \rho^{2k} \left(\mu_{m,t-k-1} \right)^2 \right) \right] = \tfrac{\rho^{2+s}}{n} \sigma_\mu^2 \left(\sum_{k=0}^{\infty} \rho^{2k} \right) \\ &= \tfrac{\rho^{2+s}}{n} \sigma_\mu^2 \tfrac{1}{1-\rho^2}. \end{aligned} \tag{49}$$

Putting together the results above, we have (for $\tau = t + s$ and $s \geq 0$)

$$\begin{gathered} \mathbf{E} \nu_{it} \nu_{j\tau} = C + D = \\ \rho^s \sigma_\alpha^2 \tfrac{1}{1-\rho^2} + \iota(i,j) \sigma_\mu^2 + \kappa(s) \tfrac{1}{n} \rho^s \sigma_\mu^2 + \tfrac{\rho^{2+s}}{n} \sigma_\mu^2 \tfrac{1}{1-\rho^2} = \\ \sigma_\alpha^2 \left(\rho^s \tfrac{1}{1-\rho^2} + \left((1 - \kappa(s)) \iota(i,j) + \kappa(s) \tfrac{1}{n} \rho^s + \tfrac{\rho^{2+s}}{n} \tfrac{1}{1-\rho^2} \right) \tfrac{\sigma_\mu^2}{\sigma_\alpha^2} \right). \end{gathered} \tag{50}$$

Using the definition $\lambda = \frac{\sigma_\mu^2}{\sigma_\alpha^2}$ and dividing by b^2 produces equation 15.

C.4.2 The structure of the covariance matrix

To write the code needed to implement the maximum liklihood estimation, it is necessary to decide how to organize the data; this organization determines the appearance of the covariance matrix. With n regions and T periods, denote $\mathbf{e}$ as the $nT \times 1$ column vector of observed emission. I organize the data so that, for example the first n elements correspond to emissions in the n regions in the first period; elements $n+1, n+2...2n$ correspond to the emissions in the n regions in the second period, and so on. The $n \times 1$ vector e_t consists of the emissions for the n regions in period t, and $\mathbf{e} = (e_1', e_2'...e_T')'$ is the column vector of observed emissions. I then write the stacked system as

$$\mathbf{e} = \mathbf{X}\beta + \mathbf{v} \text{ with } \mathbf{E}(\mathbf{v}\mathbf{v}') = \sigma^2 \mathbf{V}, \tag{51}$$

where the matrix $\mathbf{X}$ consists of the region dummies, time trends, and any other explanatory variables. (There are none of those in my application, but it would be straightforward to include them). I include a region dummy for each region: there is no "dummy variable trap" in this model, because each region receives an idiosyncratic shock, and there is also the aggregate shock. If $\sigma_\mu^2 = 0$, then $\lambda = 0$; in that case there would be a dummy variable trap, and it would then be necessary to drop one region. The parameter vector β contains the region-specific fixed effects and the coefficients of the time trends. I define $\sigma^2 = \frac{\sigma_\alpha^2}{b^2}$. It is not possible to recover the scaling factor b^2.

The upper triangular part of the covariance matrix is

$$\mathbf{V} = \frac{1}{\sigma^2}\,\mathbf{E}\begin{bmatrix} \mathbf{v}_1\mathbf{v}_1' & \mathbf{v}_1\mathbf{v}_2' & \mathbf{v}_1\mathbf{v}_3' & \cdots\cdots & \cdots\cdots & \mathbf{v}_1\mathbf{v}_T' \\ & \mathbf{v}_2\mathbf{v}_2' & \mathbf{v}_2\mathbf{v}_3' & \cdots\cdots & \cdots\cdots & \mathbf{v}_2\mathbf{v}_T' \\ & & \mathbf{v}_3\mathbf{v}_3' & \cdots\cdots & \cdots\cdots & \mathbf{v}_3\mathbf{v}_T' \\ & & & \vdots & \vdots & \vdots \\ & & & & \vdots & \vdots \\ & & & & & \mathbf{v}_T\mathbf{v}_T' \end{bmatrix} \tag{52}$$

Each of the blocks $\mathbf{E}\mathbf{v}_t\mathbf{v}_{t+s}'$ has a simple structure. Denote I_n as the n-dimensional identity matrix and denote J as the $n \times n$ matrix consisting entirely of 1's. Using equation 15) we obtain

$$\begin{gathered} \text{for } s = 0 \\ \tfrac{1}{\sigma^2}\mathbf{E}\mathbf{v}_t\mathbf{v}_{t+s} = \left(\tfrac{1}{1-\rho^2} + \tfrac{\rho^2}{n}\tfrac{1}{1-\rho^2}\lambda\right) J + \lambda I \end{gathered} \tag{53}$$

$$\begin{gathered} \text{For } s > 0 \\ \tfrac{1}{\sigma^2}\left(\rho^s\tfrac{1}{1-\rho^2} + \left(\tfrac{\rho^s}{n} + \tfrac{\rho^{2+s}}{n}\tfrac{1}{1-\rho^2}\right)\lambda\right) J. \end{gathered} \tag{54}$$

C.4.3 The maximum liklihood estimation

The formulae for the liklihood function are taken from **greene00** pages 470-471, but I use somewhat different notation. Define $\Gamma = \mathbf{V}^{-1}$. The matrix $\mathbf{V}$ is a function of ρ and λ. With $N \equiv nT$ observations, $\mathbf{v} = \mathbf{e} - \mathbf{X}\beta$, and $\mathbf{E}\mathbf{v}\mathbf{v}' = \sigma^2\mathbf{V}$, the log likelihood function is

$$\begin{aligned} \ln L &= -\tfrac{N}{2}\left[\ln\left(2\pi + \ln\sigma^2\right)\right] - \tfrac{1}{2\sigma^2}\mathbf{v}'\mathbf{V}^{-1}\mathbf{v} + \tfrac{1}{2}\ln\left|\mathbf{V}^{-1}\right| \\ &= -\tfrac{N}{2}\left[\ln\left(2\pi + \ln\sigma^2\right)\right] - \tfrac{1}{2\sigma^2}\mathbf{v}'\mathbf{V}^{-1}\mathbf{v} - \tfrac{1}{2}\ln\left|\mathbf{V}\right|. \end{aligned} \tag{55}$$

The last equality uses $\ln|\mathbf{V}^{-1}| = -\ln|\mathbf{V}|$.

Using equation 15 I have the derivatives

$$\frac{d\left(\rho^s \frac{1}{1-\rho^2}+\left(1-\kappa\iota+\kappa\frac{\rho^s}{n}+\frac{\rho^{2+s}}{n}\frac{1}{1-\rho^2}\right)\lambda\right)}{d\rho} = \tag{56}$$

$$\frac{1}{n}\frac{\rho^{s-1}}{(\rho^2-1)^2}\left(s\kappa+2\rho^2+s\rho^2-s\rho^4-2s\kappa\rho^2+s\kappa\rho^4\right)\lambda+\frac{1}{n}\frac{\rho^{s-1}}{(\rho^2-1)^2}\left(2n\rho^2+ns-ns\rho^2\right)$$

and

$$\frac{d\left(\frac{d\left(\rho^s \frac{1}{1-\rho^2}+\left(1-\kappa\iota+\kappa\frac{\rho^s}{n}+\frac{\rho^{2+s}}{n}\frac{1}{1-\rho^2}\right)\lambda\right)}{d\rho}\right)}{d\lambda} \tag{57}$$

$$= \left(1-\kappa x+\kappa\frac{\rho^s}{n}+\frac{\rho^{2+s}}{n}\frac{1}{1-\rho^2}\right),$$

where $\iota = \iota(i,j)$ and $\kappa = \kappa(s)$. With these derivatives and the matrix structure from Section C.4.2 we have $\frac{d\mathbf{V}}{d\rho}$ and $\frac{d\mathbf{V}}{d\lambda}$.

We estimate the model by estimating β and σ^2 conditional on $\mathbf{V}$ (i.e. conditional on ρ and λ), substituting those estimates into the liklihood function to obtain the concentrated liklihood function, and then maximizing that function with respect to ρ and λ.

Conditional on $\mathbf{V}$, (i.e. on λ and ρ) the first order conditions for β and σ^2 are

$$\begin{aligned}&\mathbf{X}'\Gamma(\mathbf{e}-\mathbf{X}\beta)=0\Rightarrow\\ &\tilde\beta=\left(\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{e}\end{aligned}\tag{58}$$

$$\begin{aligned}&-\tfrac{4nT}{2\sigma^2}+\tfrac{1}{2\sigma^4}(\mathbf{e}-\mathbf{x}\beta)'\,\Gamma(\mathbf{e}-\mathbf{x}\beta)=0\Rightarrow\\ &\tilde\sigma^2=\tfrac{1}{nT}\left(\mathbf{e}-\mathbf{X}'\tilde\beta\right)'\tilde{\mathbf{V}}^{-1}\left(\mathbf{e}-\mathbf{X}'\tilde\beta\right)\end{aligned}\tag{59}$$

Define $\tilde{\mathbf{v}} = \mathbf{e}-\mathbf{X}\tilde\beta$. The concentrated liklihood function is

$$L = \text{"}terms\text{"} - \frac{1}{2\sigma^2}\tilde{\mathbf{v}}'\mathbf{\Gamma}\tilde{\mathbf{v}}+\frac{1}{2}\ln|\Gamma| \tag{60}$$

where "terms" are independent of Γ. The rule for the derivative of the inverse of a matrix is:

$$\frac{dA^{-1}}{d\rho} = -A^{-1}\frac{dA}{d\rho}A^{-1}. \tag{61}$$

Consider the FOC for ρ. I have

$$\frac{\partial\tilde{\mathbf{v}}'\mathbf{\Gamma}\tilde{\mathbf{v}}}{\partial\rho} = \frac{\partial\tilde{\mathbf{v}}'\mathbf{V}^{-1}\tilde{\mathbf{v}}}{\partial\rho} = -\tilde{\mathbf{v}}'\mathbf{V}^{-1}\frac{d\mathbf{V}}{d\rho}\mathbf{V}^{-1}\tilde{\mathbf{v}} \tag{62}$$

I use

$$\partial\left(\ln|\Gamma|\right) = Tr\left(\Gamma^{-1}\partial\Gamma\right) \Rightarrow \frac{\partial\left(\ln|\Gamma|\right)}{\partial\rho} = Tr\left(\Gamma^{-1}\frac{\partial\Gamma}{\partial\rho}\right) \tag{63}$$

$$= -Tr\left(\mathbf{V}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\rho}\mathbf{V}^{-1}\right) = -Tr\left(\frac{\partial\mathbf{V}}{\partial\rho}\mathbf{V}^{-1}\right) \tag{64}$$

Using these two equations I have the FOC for ρ:

$$\frac{d\ln L}{d\rho} = \frac{1}{2\sigma^2}\tilde{\mathbf{v}}'\mathbf{V}^{-1}\frac{d\mathbf{V}}{d\rho}\mathbf{V}^{-1}\tilde{\mathbf{v}} - Tr\left(\frac{d\mathbf{V}}{d\rho}\mathbf{V}^{-1}\right) = 0 \tag{65}$$

I also have the FOC for λ

$$\frac{d\ln L}{d\lambda} = \frac{1}{2\sigma^2}\tilde{\mathbf{v}}'\mathbf{V}^{-1}\frac{d\mathbf{V}}{d\lambda}\mathbf{V}^{-1}\tilde{\mathbf{v}} - Tr\left(\frac{d\mathbf{V}}{d\lambda}\mathbf{V}^{-1}\right) = 0 \tag{66}$$

However (at least with for the data I use) ML estimation always produces the smallest feasible value of λ (e.g. $\lambda = 0$) for any estimate of ρ. To understand why this happens, use the first order condition 58 to write

$$\frac{1}{2\hat{\sigma}^2}\tilde{\mathbf{v}}'\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{v}} = \frac{N}{2\tilde{\mathbf{v}}'\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{v}}}\tilde{\mathbf{v}}'\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{v}} = \frac{N}{2}. \tag{67}$$

Subsituting this expression into equation 55 produces an alternative version of the concentrated liklihood function,

$$\ln L = -\frac{N}{2}\left[1+\ln\left(2\pi\right)\right] - \frac{N}{2}\left[\ln\hat{\sigma}^2\right] - \frac{1}{2}\ln\left|\hat{\mathbf{V}}\right|, \tag{68}$$

with

$$\tilde{\sigma}^2 = \frac{1}{nT}\left(\mathbf{e}-\mathbf{X}'\tilde{\beta}\right)'\tilde{\mathbf{V}}^{-1}\left(\mathbf{e}-\mathbf{X}'\tilde{\beta}\right) \tag{69}$$

For $\lambda = 0$, the matrix $\mathbf{V}$ (and also $\tilde{\mathbf{V}}$) is singular, as is apparent from equations 52 – 54. Given continuity $-\frac{1}{2}\ln\left|\tilde{\mathbf{V}}\right| \to \infty$ as $\lambda \to 0$. Therefore, a sufficient condition for $\lambda = 0$ to maximize $\ln L$ is that $\ln\tilde{\sigma}^2$ remains bounded as λ becomes small. This condition is met, at least with the data that I use.

C.5 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 15 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 $\sigma_\alpha^2 \geq 0$. Greene (pg 571) notes a similar issue in estimating the simpler random effects mode. For the carbon emissions data, the point estimates obtained using the alternative here does not satisfy this inequality.

Using the first equation in system 13,

$$\frac{1}{n}\sum_i \nu_{it} = \frac{1}{n}\sum_i \left(\nu_t + \mu_{it} - \frac{\sum_j \mu_{jt}}{n}\right) = \nu_t. \tag{70}$$

Using equation 14, average emissions at t equals

$$\bar{e}_t = \frac{1}{n}\sum_i e_{it} = \frac{1}{b}\left(\frac{\sum_i b_{0i}}{n} + h(t) + \frac{\sum_i \nu_{it}}{n}\right) = \frac{1}{b}\left(B_0 + h(t) + \nu_t\right), \tag{71}$$

where the second equality uses equations 6 and 70. Using $\nu_t = \rho\nu_{t-1} + \theta_t$ with $\theta_t \sim (0, \sigma^2)$, we can estimate equation 71 using GLS or by transforming the data.

With the later method, we Lag equation 71 and multiply by ρ to obtain

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

Adding this expression to the right side of equation 71 produces the regression equation

$$\begin{gathered}\bar{e}_t = \rho\bar{e}_{t-1} + \tfrac{1}{b}\left[(1-\rho)B_0 + h(t) - \rho h(t-1)\right] + \omega_t \\ \text{with } \omega_t \equiv \tfrac{\theta_t}{b} \sim iid\left(0, \tfrac{\sigma^2}{b^2}\right) \\ \text{and } \theta_t \equiv \alpha_t + \tfrac{\sum_j \mu_{jt}}{n} \Rightarrow \sigma^2 = \sigma_\alpha^2 + \tfrac{\sigma_\mu^2}{n}\end{gathered} \tag{72}$$

[I dropped the scaling factor $\frac{1}{B}$. It will be clearer to retain this factor.] 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 71. Using $\nu_t = \rho\nu_{t-1} + \theta_t$,

$$\begin{aligned} \nu_t = \sum_{k=0}^{\infty} \rho^k \theta_{t-k} \Rightarrow \mathbf{E}\nu_t\nu_{t-s} &= \mathbf{E}\left(\sum_{k=0}^{\infty} \rho^k \theta_{t-k}\right)\left(\sum_{k=0}^{\infty} \rho^k \theta_{t-s-k}\right) \\ &= \mathbf{E}\left(\sum_{k=0}^{s-1} \rho^k \theta_{t-k} + \sum_{k=s}^{\infty} \rho^k \theta_{t-k}\right)\left(\sum_{k=0}^{\infty} \rho^k \theta_{t-s-k}\right) \\ &= \mathbf{E}\left(\sum_{k=0}^{s-1} \rho^k \theta_{t-k} + \rho^s \sum_{k=0}^{\infty} \rho^k \theta_{t-k-s}\right)\left(\sum_{k=0}^{\infty} \rho^k \theta_{t-s-k}\right) \\ &= \rho^s \sigma^2 \sum_{k=0}^{\infty} \rho^{2k} = \rho^s \sigma^2 \frac{1}{1-\rho^2} \end{aligned} \tag{73}$$

The third line uses the definition $m = k - s$ to write

$$\sum_{k=s}^{\infty} \rho^k \theta_{t-k} = \sum_{m=0}^{\infty} \rho^{m+s} \theta_{t-m-s} = \rho^s \sum_{k=0}^{\infty} \rho^{k+s} \theta_{t-k-s}). \tag{74}$$

GLS produces more efficient point estimates. The estimation of

$$\begin{gathered} \bar{e}_t = \frac{1}{b}\left(B_0 + h(t) + \nu_t\right) \\ \text{with} \\ \mathbf{E}\nu_t\nu_{t-s} = \sigma^2 \frac{\rho^s}{1-\rho^2} \end{gathered} \tag{75}$$

produces estimates of $\frac{B_0}{b}$, $\frac{h(t)}{b}$, ρ, and $\sigma^2 = \sigma_\alpha^2 + \frac{\sigma_\mu^2}{n}$

I now consider the estimation of the region-specific effects and of σ_μ^2. Because $\sum_{i=1}^{n}\left(e_{i,t} - \bar{e}_t\right) = 0$, I drop one region, the n'th, to avoid a singular covariance matrix when estimating the parameters of equation 78, $\frac{b_{0i} - B_0}{b}$, $i = 1, 2..n$ and $\frac{\sigma_\mu^2}{nb^2}$. Due to the cross-regional correlation of the errors, $\epsilon_{it} \equiv \frac{1}{b}\left(\mu_{it} - \frac{1}{n}\sum_i \mu_{it}\right)$, I estimate this equation using GLS. The formulae for these GLS estimators 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;
- $\mathbf{y}_t = \left(e_{1,t} - \bar{e}_t, e_{2,t} - \bar{e}_t, ...e_{n-1,t} - \bar{e}_t\right)'$ is the vector of deviations of region i's emissions from the average emissions in period t and $\mathbf{y} = \left(\mathbf{y}_1', \mathbf{y}_2', ...\mathbf{y}_T'\right)'$ is the stacked $(n-1)T$ column vector of observations;
- from equation 78, $\mathbf{E}\left(\epsilon_{1t}, \epsilon_{2t}, ...\epsilon_{n-1,t}\right)'\left(\epsilon_{1t}, \epsilon_{2t}, ...\epsilon_{n-1,t}\right) = \frac{\sigma_\mu^2}{nb^2}\Omega$, with the $n-1$ by $n-1$

matrix

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

Remark 1. *The GLS estimator of $\frac{b_{0i}-B_0}{b}$ is*

$$\frac{\widehat{b_{0i}-B_0}}{b} = \frac{e_i^{AG} - \frac{e^{AG}}{n}}{T} \tag{76}$$

and the estimate of the variance is

$$\frac{\widehat{\sigma_\mu^2}}{nb^2} = \frac{\mathbf{y}'\left[\left(I_T - \frac{1}{T}\mathbf{ss}'\right) \otimes \Omega^{-1}\right]\mathbf{y}}{(n-1)\,T - n} \tag{77}$$

The expressions for $\bar{e}_t$ and e_{it} imply[19]

$$e_{it} - \bar{e}_t = \frac{(b_{0i} - B_0)}{b} + \epsilon_{it} \text{ with } \epsilon_{it} \equiv \frac{1}{b}\left(\mu_{it} - \frac{\sum_i \mu_{it}}{n}\right), \; i = 1, 2...n \tag{78}$$

Estimating the parameters of equation 78 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 78, $\mathbf{E}\left(\epsilon_{1t}, \epsilon_{2t}, ...\epsilon_{n-1,t}\right)'\left(\epsilon_{1t}, \epsilon_{2t}, ...\epsilon_{n-1,t}\right) = \frac{\sigma_\mu^2}{nb^2}\Omega$, with the $n-1$ by $n-1$

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

matrix

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

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

$$\widehat{\frac{b_{0i}-B_0}{b}} = \frac{e_i^{AG} - \frac{e^{AG}}{n}}{T} \tag{79}$$

and the estimate of the variance is

$$\widehat{\frac{\sigma_\mu^2}{nb^2}} = \frac{\mathbf{y}'\left[\left(I_T - \frac{1}{T}\mathbf{s}\mathbf{s}'\right) \otimes \Omega^{-1}\right]\mathbf{y}}{(n-1)\,T - n} \tag{80}$$

I now provide the details. Estimation of the parameters of equation 78 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 $\tilde{\epsilon}_t = (\epsilon_{1,t}, \epsilon_{2,t}...\epsilon_{n,t})'$, 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}, $\mathbf{E}\tilde{\epsilon}_t\tilde{\epsilon}_t' = \frac{\sigma_\mu^2}{nb^2}\tilde{\Omega}$. 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 $\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t}...\epsilon_{n-1,t})'$, the matrix obtained by dropping the last row and column of $\tilde{\Omega}$. The tilda on $\tilde{\epsilon}_t$ distinguishes the n-dimensional column vector from ϵ_t, the $n-1$-dimensional vector.

Proof. (Remark 1) 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 $f_i = \frac{b_{0i} - B_0}{b}$; $\epsilon = (\epsilon_1', \epsilon_2', ...\epsilon_T')'$ is the column vector of errors of the $n-1$ included regions.

With this notation, the regression equation 78 becomes

$$\mathbf{y} = X\mathbf{f} + \epsilon \text{ with } \mathbf{E}\epsilon = \mathbf{0} \text{ and } \mathbf{E}\epsilon\epsilon' = (I_T \otimes \Omega) \frac{\sigma_\mu^2}{nb^2} \tag{81}$$

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 $(I \otimes V)(I \otimes V)'$. Pre-multiply the regression 81 by $(I_T \otimes V)^{-1} = (I_T \otimes V^{-1})$ to obtain

$$\left(I_T \otimes V^{-1}\right)\mathbf{y} = \left(I_T \otimes V^{-1}\right)X\mathbf{f} + \left(I_T \otimes V^{-1}\right)\epsilon. \tag{82}$$

The OLS estimator for the transformed system 82 (equivalent to the GLS estimator for the untransformed system 81) 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,

$$\hat{\mathbf{f}} = \frac{\Omega}{T}\left(\mathbf{s}' \otimes \Omega^{-1}\right)\mathbf{y} = \frac{1}{T}\left(\mathbf{s}' \otimes I_{n-1}\right)\mathbf{y} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{y}_t, \tag{83}$$

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 6, $\sum_{j=1}^{n} f_j = 0$, 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 79.

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 82 and 83, the vector of residual is

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

Therefore, sum of squared residuals, SSR, equals

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

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

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

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

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

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

$$\frac{\sigma_\mu^2}{nb^2} = \frac{\mathbf{y}' \left[\left(I_T - \frac{1}{T}\mathbf{ss}' \right) \otimes \Omega^{-1} \right] \mathbf{y}}{(n-1)\,T - n} \tag{84}$$

□

C.6 A third estimation approach

The third approach begins by using the averages across regions to estimate

$$\begin{gathered} \bar{e}_t = \tfrac{1}{b}\left(B_0 + h\left(t\right) + \nu_t\right) \\ \text{with} \\ \mathbf{E}\nu_t\nu_{t-s} = \sigma^2 \tfrac{\rho^s}{1-\rho^2} \end{gathered} \tag{85}$$

The formula for the covariance uses

$$\begin{gathered} \nu_t = \rho\nu_{t-1} + \theta_t = \textstyle\sum_{k=0}^{\infty} \rho^k \theta_{t-k} = \sum_{k=0}^{s-1} \rho^k \theta_{t-k} + \sum_{k=s}^{\infty} \rho^k \theta_{t-k} = \\ (\text{define } m = k - s)\ \textstyle\sum_{k=0}^{s-1} \rho^k \theta_{t-k} + \sum_{m=0}^{\infty} \rho^{m+s} \theta_{t-m-s} = \\ \textstyle\sum_{k=0}^{s-1} \rho^k \theta_{t-k} + \rho^s \sum_{k=0}^{\infty} \rho^k \theta_{t-k-s} \text{ and} \\ \nu_{t-s} = \textstyle\sum_{k=s}^{\infty} \rho^k \theta_{t-s-k} \Rightarrow \\ \mathbf{E}\nu_t\nu_{t-s} = \mathbf{E}\left(\textstyle\sum_{k=0}^{s-1} \rho^k \theta_{t-k} + \rho^s \sum_{k=0}^{\infty} \rho^k \theta_{t-k-s}\right)\left(\sum_{k=s}^{\infty} \rho^k \theta_{t-s-k}\right) = \\ \mathbf{E}\left(\rho^s \textstyle\sum_{k=0}^{\infty} \rho^k \theta_{t-k-s}\right)\left(\sum_{k=s}^{\infty} \rho^k \theta_{t-s-k}\right) = \rho^s \sigma_\theta^2 \sum_{k=0}^{\infty} \rho^{2k} = \rho^s \sigma_\theta^2 \frac{1}{1-\rho^2}. \end{gathered} \tag{86}$$

Estimate this model using GLS to obtain estimates of $\frac{h(t)}{b}$ and ρ. The estimation also produces an estimate of $\frac{\sigma_\theta^2}{b}$, but I don't expect to use that.

Step 1 Estimate 85 using GLS to produce estimates of $\frac{h(t)}{b}$ and ρ.

Step 2 Detrend the data and substract regional averages. Detrending gives $z_{it} \equiv e_{it} - \frac{\hat{h}(t)}{b} = \frac{1}{b}\left(b_{0i} + h\left(t\right) + \nu_{i,t}\right) - \frac{\hat{h}(t)}{b}$. Of course the trend term depends on the data but I am going to ignore this and proceed as if $\frac{1}{b}h\left(t\right) - \frac{\hat{h}(t)}{b} = 0$, so $z_{it} \equiv e_{it} - \frac{\hat{h}(t)}{b} = \frac{1}{b}\left(b_{0i} + \nu_{i,t}\right)$. Now average over time to obtain $\bar{z}_i = \frac{1}{T}\sum_t z_{it} = \frac{b_{0i}}{b} + \frac{1}{bT}\sum_t \nu_{it}$. Subtract the mean to obtain $y_{it} \equiv z_{it} - \bar{z}_i = \frac{1}{b}\left(\nu_{i,t} - \frac{1}{T}\sum_{s=1}^{T} \nu_{is}\right)$. I drop the last element to avoid a singlular covariance matrix, so I have $n\left(T-1\right)$ zero mean observatons. The index s in the summation runs from 1 to T but the index t in ν_{it} runs from 1 to $T-1$.

Step 3 Find the covariance matrix for y in terms of σ_α^2, σ_μ^2 and ρ. I have

$$\begin{aligned}\mathbf{E}\left(\nu_{it}-\tfrac{1}{T}\textstyle\sum_t \nu_{it}\right)\left(\nu_{j\tau}-\tfrac{1}{T}\textstyle\sum_t \nu_{jt}\right) &= \mathbf{E}\left(\nu_{it}\right)\left(\nu_{j\tau}\right)-\mathbf{E}\left(\nu_{it}\right)\left(\tfrac{1}{T}\textstyle\sum_t \nu_{jt}\right) \\ &-\mathbf{E}\left(\tfrac{1}{T}\textstyle\sum_t \nu_{it}\right)\left(\nu_{j\tau}\right)+\mathbf{E}\left(\tfrac{1}{T}\textstyle\sum_t \nu_{it}\right)\left(\tfrac{1}{T}\textstyle\sum_t \nu_{jt}\right).\end{aligned} \tag{87}$$

I have already computed the first term on the right side, given by equation 50 (using $\tau = t+s$):

$$\begin{gathered}F_0 \equiv \mathbf{E}\left(\nu_{it}\right)\left(\nu_{j\tau}\right)= \\ \tfrac{\rho^s}{1-\rho^2}\sigma_\alpha^2+\left(\iota\left(i,j\right)+\kappa\left(s\right)\tfrac{1}{n}\rho^s+\tfrac{\rho^{2+s}}{n}\tfrac{1}{1-\rho^2}\right)\sigma_\mu^2\end{gathered} \tag{88}$$

I need to compute

$$F_1 \equiv \mathbf{E}\left(\nu_{it}\right)\left(\frac{1}{T}\sum_t \nu_{jt}\right) \tag{89}$$

$$F_2 \equiv \mathbf{E}\left(\frac{1}{T}\sum_t \nu_{it}\right)\left(\frac{1}{T}\sum_t \nu_{jt}\right) \tag{90}$$

(Note that $\mathbf{E}\left(\nu_{it}\right)\left(\frac{1}{T}\sum_t \nu_{jt}\right)=\mathbf{E}\left(\nu_{j\tau}\right)\left(\frac{1}{T}\sum_t \nu_{it}\right)$ by symmetry.)

Using equation 42 I have

$$\nu_{it}=\left(\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)+\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_i \mu_{i,t-k-1}}{n}\right)+\mu_{it}\right) \tag{91}$$

Therefore

$$\left(\frac{1}{T}\sum_{s=1}^{T}\nu_{is}\right)=\frac{1}{T}\sum_{s=1}^{T}\left(\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right)+\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_{i=1}^{n}\mu_{i,s-k-1}}{n}\right)+\mu_{is}\right). \tag{92}$$

Notice that the i index shows up in two places on the right side. With these results, and with a view to calculating F_1 I use

$$\begin{gathered}F_1 \equiv \left(\nu_{i\tau}\right)\left(\tfrac{1}{T}\textstyle\sum_t \nu_{jt}\right)= \\ \left(\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)+\rho\textstyle\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_i \mu_{i,t-k-1}}{n}\right)+\mu_{it}\right)\tfrac{1}{T}\left[\textstyle\sum_{s=1}^{T}\left(\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right)+\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_{i=1}^{n}\mu_{j,s-k-1}}{n}\right)+\right.\right.\end{gathered} \tag{93}$$

Consider the expectation of each product:

$$F_{11} = \frac{1}{T}\mathbf{E}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left[\sum_{s=1}^{T}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right)\right] \tag{94}$$

$$F_{12} = \frac{1}{T}\mathbf{E}\left[\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_i \mu_{i,t-k-1}}{n}\right)\right)+\mu_{it}\right]\left[\sum_{s=1}^{T}\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_{i=1}^{n}\mu_{j,s-k-1}}{n}\right)\right)\right] \tag{95}$$

$$F_{13} = \frac{1}{T}\mathbf{E}\left[\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_i \mu_{i,t-k-1}}{n}\right)\right)+\mu_{it}\right]\left[\sum_{s=1}^{T}\mu_{js}\right] \tag{96}$$

The first product is

$$F_{11} = \frac{1}{T}\mathbf{E}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left[\sum_{s=1}^{T}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right)\right] \tag{97}$$

Use equation 44 to write

$$\text{for } s \geq t\text{: } \mathbf{E}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right) = \rho^{s-t}\sigma_\alpha^2\frac{1}{1-\rho^2} \tag{98}$$

$$\text{for } s < t\text{: } \mathbf{E}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right) = \rho^{t-s}\sigma_\alpha^2\frac{1}{1-\rho^2}. \tag{99}$$

I derived the first of these two equations for $s \geq t$ in equation 44. To derive the second with $s < t$ I merely switch the s and t indices. The point is that the exponent on ρ^{s-t} should always be understood as $\rho^{|s-t|}$. But it is awkward to write the absolute value in the exponent so I write out the two cases.So I have

$$\begin{aligned}
F_{11} &= \tfrac{1}{T}\mathbf{E}\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left[\textstyle\sum_{s=1}^{T}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right)\right]\\
&= \tfrac{1}{T}\mathbf{E}\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left[\textstyle\sum_{s=1}^{t-1}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right)+\sum_{s=t}^{T}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right)\right]\\
&= \tfrac{1}{T}\mathbf{E}\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left[\textstyle\sum_{s=1}^{t-1}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right)\right]+\tfrac{1}{T}\mathbf{E}\left(\textstyle\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left[\textstyle\sum_{s=t}^{T}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{s-k}\right)\right]\\
&= \tfrac{1}{T}\sigma_\alpha^2\tfrac{1}{1-\rho^2}\left[\textstyle\sum_{s=1}^{t-1}\rho^{t-s}+\sum_{s=t}^{T}\rho^{s-t}\right]\\
&= \tfrac{1}{T}\sigma_\alpha^2\tfrac{1}{1-\rho^2}\left(\tfrac{1+\rho-\rho^t-\rho^{T+1-t}}{1-\rho}\right).
\end{aligned} \tag{100}$$

(Use $\sum_{s=1}^{t-1}\rho^{t-s} = \frac{\rho-\rho^t}{1-\rho}$ is true and $\sum_{s=t}^{T}\rho^{s-t} = \frac{1-\rho^{T+1-t}}{1-\rho}$ is true. $\frac{\rho-\rho^t}{1-\rho}+\frac{1-\rho^{T+1-t}}{1-\rho} = \frac{1+\rho-\rho^t-\rho^{T+1-t}}{1-\rho}$ is true. Alternatively use $\sum_{s=1}^{t-1}\rho^{t-s}+\sum_{s=t}^{T}\rho^{s-t} = \frac{\rho-\rho^t-\rho^{T+1-t}+1}{1-\rho}$)

In calculating F_{12} I initially thought that I had to distinguish between the cases $i = j$ and $i \neq j$, but on reflection I see that is not the case. Consider

$$F_{12} = \frac{1}{T}\mathbf{E}\left[\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\frac{\sum_{i=1}^{n} \mu_{i,t-k-1}}{n}\right)\right) + \mu_{it}\right]\left[\sum_{s=1}^{T}\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\frac{\sum_{i=1}^{n} \mu_{j,s-k-1}}{n}\right)\right)\right] \quad (101)$$

I can change the index in the sum $\sum_{i=1}^{n} \mu_{i,t-k-1}$ to write it as $\sum_{p=1}^{n} \mu_{p,t-k-1}$, so that this sum does not depend on i. This is legitimate because the sum is over all regions, so the sum does not depend on the particular region. Thus I can write

$$F_{12} = \frac{1}{T}\mathbf{E}\left[\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\frac{\sum_{p=1}^{n} \mu_{p,t-k-1}}{n}\right)\right) + \mu_{it}\right]\left[\sum_{s=1}^{T}\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\frac{\sum_{q=1}^{n} \mu_{q,s-k-1}}{n}\right)\right)\right] \quad (102)$$

The only term that involves i is in the first term, but for all i, the expectation

$$\mathbf{E}\left[\mu_{it} \sum_{s=1}^{T}\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\frac{\sum_{q=1}^{n} \mu_{q,s-k-1}}{n}\right)\right)\right] = \mathbf{E}\left[\mu_{it} \sum_{s=1}^{T}\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\frac{\mu_{i,s-k-1}}{n}\right)\right)\right], \quad (103)$$

which is the same for all i. This expression is independent of i because every product involves a sum over regions; therefore, for every i, the expectation of $\mu_{i.t}$ and the sum returns only one nonzero element.

For F_{12} I have

$$\begin{gathered} F_{12} = \tfrac{1}{T}\mathbf{E}\left[\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\tfrac{\sum_i \mu_{i,t-k-1}}{n}\right)\right) + \mu_{it}\right]\left[\sum_{s=1}^{T}\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\tfrac{\sum_{i=1}^{n} \mu_{i,s-k-1}}{n}\right)\right)\right] \\ = F_{12a} + F_{12b} \\ \text{with} \\ F_{12a} = \tfrac{1}{T}\mathbf{E}\left[\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\tfrac{\sum_i \mu_{i,t-k-1}}{n}\right)\right)\right]\left[\sum_{s=1}^{T}\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\tfrac{\sum_{i=1}^{n} \mu_{i,s-k-1}}{n}\right)\right)\right] \\ F_{12b} = \tfrac{1}{T}\mathbf{E}\left[\mu_{it}\right]\left[\sum_{s=1}^{T}\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\tfrac{\sum_{i=1}^{n} \mu_{i,s-k-1}}{n}\right)\right)\right] \end{gathered} \quad (104)$$

Consider F_{12a}. For a particular s I have

$$\begin{gathered} \tfrac{1}{T}\mathbf{E}\left[\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\tfrac{\sum_i \mu_{i,t-k-1}}{n}\right)\right)\right]\left[\left(\rho \sum_{k=0}^{\infty} \rho^k \left(\tfrac{\sum_{i=1}^{n} \mu_{i,s-k-1}}{n}\right)\right)\right] = \\ \tfrac{1}{T}\left(\tfrac{\rho}{n}\right)^2 \mathbf{E}\left[\left(\sum_{k=0}^{\infty}\left[\rho^k \left(\sum_i \mu_{i,t-k-1}\right)\right]\right)\right]\left[\left(\sum_{k=0}^{\infty} \rho^k \left(\sum_{i=1}^{n} \mu_{i,s-k-1}\right)\right)\right] \end{gathered} \quad (105)$$

For a particular i this equals

$$\frac{1}{T}\left(\frac{\rho}{n}\right)^2 \mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\left(\mu_{i,t-k-1}\right)\right)\right]\left[\left(\sum_{k=0}^{\infty}\rho^k\left(\mu_{i,s-k-1}\right)\right)\right] \tag{106}$$

I can use the rule in equations 98 and 99 to write

$$\text{for } s \;\geq\; t\text{: } \mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\left(\mu_{i,t-k-1}\right)\right)\right]\left[\left(\sum_{k=0}^{\infty}\rho^k\left(\mu_{i,s-k-1}\right)\right)\right] = \rho^{s-t}\sigma_\mu^2\frac{1}{1-\rho^2} \tag{107}$$

$$\text{for } s \;<\; t\text{: } \mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\left(\mu_{i,t-k-1}\right)\right)\right]\left[\left(\sum_{k=0}^{\infty}\rho^k\left(\mu_{i,s-k-1}\right)\right)\right] = \rho^{t-s}\sigma_\mu^2\frac{1}{1-\rho^2} \tag{108}$$

For the particular s I have n of these objects, so for a particular $s \geq t$ I have I have

$$\begin{gathered}\tfrac{1}{T}\mathbf{E}\left[\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\tfrac{\sum_i \mu_{i,t-k-1}}{n}\right)\right)\right]\left[\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\tfrac{\sum_{i=1}^n \mu_{i,s-k-1}}{n}\right)\right)\right] = \\ \tfrac{1}{T}\left(\tfrac{\rho}{n}\right)^2 n\rho^{s-t}\sigma_\mu^2\tfrac{1}{1-\rho^2} = \tfrac{1}{T}\left(\tfrac{1}{n}\right)\rho^2\rho^{s-t}\sigma_\mu^2\tfrac{1}{1-\rho^2}\end{gathered} \tag{109}$$

And for $s < t$ I have

$$\begin{gathered}\tfrac{1}{T}\mathbf{E}\left[\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\tfrac{\sum_i \mu_{i,t-k-1}}{n}\right)\right)\right]\left[\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\tfrac{\sum_{i=1}^n \mu_{i,s-k-1}}{n}\right)\right)\right] = \\ \tfrac{1}{T}\left(\tfrac{1}{n}\right)\rho^2\rho^{t-s}\sigma_\mu^2\tfrac{1}{1-\rho^2}.\end{gathered} \tag{110}$$

Summing over s I have

$$\begin{aligned}F_{12a} &= \tfrac{1}{T}\mathbf{E}\left[\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\tfrac{\sum_i \mu_{i,t-k-1}}{n}\right)\right)\right]\left[\sum_{s=1}^T\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\tfrac{\sum_{i=1}^n \mu_{i,s-k-1}}{n}\right)\right)\right] \\ &= \tfrac{1}{T}\left(\tfrac{1}{n}\right)\rho^2\sigma_\mu^2\tfrac{1}{1-\rho^2}\left[\sum_{s=1}^{t-1}\rho^{t-s}+\sum_{s=t}^{T}\rho^{s-t}\right] = \tfrac{1}{T}\left(\tfrac{1}{n}\right)\rho^2\sigma_\mu^2\tfrac{1}{1-\rho^2}\tfrac{1+\rho-\rho^t-\rho^{T+1-t}}{1-\rho}\end{aligned} \tag{111}$$

(Use $\sum_{s=1}^{t-1}\rho^{t-s}+\sum_{s=t}^{T}\rho^{s-t} = \frac{1+\rho-\rho^t-\rho^{T+1-t}}{1-\rho}$ is true)

Now calculate F_{12b}

$$F_{12b} = \frac{1}{T}\mathbf{E}\left[\mu_{it}\right]\left[\sum_{s=1}^{T}\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_{i=1}^n \mu_{i,s-k-1}}{n}\right)\right)\right] \tag{112}$$

For a particular s this equals

$$\frac{1}{T}\mathbf{E}\left[\mu_{it}\right]\left[\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_{i=1}^n \mu_{i,s-k-1}}{n}\right)\right)\right]. \tag{113}$$

Each term is zero except for $s-k-1=t$, i.e. for $s-t-1=k$. Therefore,

$$\frac{1}{T}\mathbf{E}\left[\mu_{it}\right]\left[\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\frac{\sum_{i=1}^{n}\mu_{i,s-k-1}}{n}\right)\right)\right]=\frac{1}{T}\left(\rho\right)\rho^{s-t-1}\sigma_{\mu}^{2}=\frac{1}{T}\rho^{s-t}\sigma_{\mu}^{2}. \tag{114}$$

Recall that k are nonnegative integers, so this term is nonzero only for $s\geq t$. Therefore

$$\begin{gathered}F_{12b}=\tfrac{1}{T}\mathbf{E}\left[\mu_{it}\right]\left[\sum_{s=1}^{T}\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\tfrac{\sum_{i=1}^{n}\mu_{i,s-k-1}}{n}\right)\right)\right]\\ =\tfrac{1}{T}\sigma_{\mu}^{2}\sum_{s=t}^{T}\rho^{s-t}=\tfrac{1}{T}\sigma_{\mu}^{2}\tfrac{1-\rho^{T+1-t}}{1-\rho}\end{gathered} \tag{115}$$

Use $\sum_{s=t}^{T}\rho^{s-t}=\frac{1-\rho^{T+1-t}}{1-\rho}$ is true.

Now use

$$\begin{gathered}F_{12}=F_{12a}+F_{12b}=\\ \tfrac{1}{T}\sigma_{\mu}^{2}\left(\left(\tfrac{1}{n}\right)\rho^{2}\tfrac{1}{1-\rho^{2}}\tfrac{1+\rho-\rho^{t}-\rho^{T+1-t}}{1-\rho}+\tfrac{1-\rho^{T+1-t}}{1-\rho}\right)\end{gathered} \tag{116}$$

($F_{12a}\left(i=j\right)+F_{12b}\left(i=j\right)=\frac{1}{T}\left(\frac{1}{n}\right)\rho^{2}\sigma_{\mu}^{2}\frac{1}{1-\rho^{2}}\frac{1+\rho-\rho^{t}-\rho^{T+1-t}}{1-\rho}+\frac{1}{T}\sigma_{\mu}^{2}\frac{1-\rho^{T+1-t}}{1-\rho}=\frac{1}{T}\sigma_{\mu}^{2}\left(\left(\frac{1}{n}\right)\rho^{2}\frac{1}{1-\rho^{2}}\frac{1+\rho-\rho^{t}-\rho^{T+1-t}}{1-\rho}+\right.$ 1
is true.)

Now calculate

$$\begin{gathered}F_{13}=\tfrac{1}{T}\mathbf{E}\left[\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\tfrac{\sum_{p=1}^{n}\mu_{p,t-k-1}}{n}\right)\right)+\mu_{it}\right]\left[\sum_{s=1}^{T}\mu_{js}\right]=\\ \tfrac{1}{T}\mathbf{E}\left[\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\tfrac{\mu_{j,t-k-1}}{n}\right)\right)+\mu_{it}\right]\left[\sum_{s=1}^{T}\mu_{js}\right]=\\ \tfrac{1}{T}\mathbf{E}\left[\left[\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\tfrac{\mu_{j,t-k-1}}{n}\right)\right)\right]\left[\sum_{s=1}^{T}\mu_{js}\right]+\mu_{it}\left[\sum_{s=1}^{T}\mu_{js}\right]\right].\end{gathered} \tag{117}$$

The first equality follows because for any s, $\mathbf{E}\mu_{p,t-k-1}\mu_{js}$ is certainly zero unless $p=j$. Notice that F_{13} does depend on the indices i,j. The second term equals

$$\frac{1}{T}\mathbf{E}\left[\mu_{it}\left[\sum_{s=1}^{T}\mu_{js}\right]\right]=\begin{cases}\frac{\sigma_{\mu}^{2}}{T}\text{ if }i=j\\ 0\text{ if }i\neq j.\end{cases} \tag{118}$$

The first term equals

$$\frac{1}{T}\mathbf{E}\left[\left[\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\frac{\mu_{j,t-k-1}}{n}\right)\right)\right]\left[\sum_{s=1}^{T}\mu_{js}\right]\right] \tag{119}$$

Consider an arbitrary s this expectation equals

$$\frac{1}{T}\mathbf{E}\left[\left[\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\frac{\mu_{j,t-k-1}}{n}\right)\right)\right]\left[\mu_{js}\right]\right]=\begin{cases}\left(\frac{1}{nT}\rho\right)\rho^{t-s-1}\sigma_{\mu}^{2}\text{ if }s\leq t-1\\ 0\text{ if }s>t.\end{cases}\tag{120}$$

(Eventually I want to write the expressions using t and $\tau = t+s$. Then I will arrange exponents to make the absolute value sign unnecessary.) For example, for $t = 1$, the expectation is zero because $s \geq 1$. Therefore

$$\frac{1}{T}\mathbf{E}\left[\left[\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\frac{\mu_{j,t-k-1}}{n}\right)\right)\right]\left[\sum_{s=1}^{T}\mu_{js}\right]\right]=\begin{cases}\frac{1}{nT}\rho\sigma_{\mu}^{2}\rho^{t-s-1}\text{ if }s\leq t-1\\ 0\text{ if }s>t\end{cases}=\begin{cases}\frac{1}{nT}\sigma_{\mu}^{2}\rho^{t-s}\text{ if }s\leq t-1\\ 0\text{ if }s>t\end{cases}\tag{121}$$

If I want to use the convention that $\tau \geq t$ and define $\tau - t = s$ I will have to change notation above.

Start here next time

****

Now calculate

$$\begin{gathered}F_{12}\left(i\neq j\right)=\\ \tfrac{1}{T}\mathbf{E}\left[\left(\rho\textstyle\sum_{k=0}^{\infty}\rho^{k}\left(\frac{\sum_{i=1}^{n}\mu_{i,t-k-1}}{n}\right)\right)+\mu_{it}\right]\left[\textstyle\sum_{s=1}^{T}\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\frac{\sum_{i=1}^{n}\mu_{j,s-k-1}}{n}\right)\right)\right]\end{gathered}\tag{122}$$

xxxxxxxxxxxxxxxxxxxxxx

Again I have to consider two cases, $s \geq t$ and $s < t$. For $s \geq t$ I have

$$\frac{1}{T}\left(\frac{\rho}{n}\right)^{2}\mathbf{E}\left[\left(\sum_{k=0}^{\infty}\left[\rho^{k}\left(\mu_{i,t-k-1}\right)\right]\right)\right]\left[\left(\sum_{k=0}^{\infty}\rho^{k}\left(\mu_{i,s-k-1}\right)\right)\right]\tag{123}$$

$$\begin{aligned}F_{12a}&=\tfrac{1}{T}\mathbf{E}\left[\left(\rho\textstyle\sum_{k=0}^{\infty}\rho^{k}\left(\frac{\sum_{i}\mu_{i,t-k-1}}{n}\right)\right)\right]\left[\textstyle\sum_{s=1}^{T}\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\frac{\sum_{i=1}^{n}\mu_{i,s-k-1}}{n}\right)\right)\right]\\&=\tfrac{1}{Tn^{2}}\mathbf{E}\left[\left(\rho\textstyle\sum_{k=0}^{\infty}\rho^{k}\left(\sum_{i}\mu_{i,t-k-1}\right)\right)\right]\left[\textstyle\sum_{s=1}^{T}\left(\rho\sum_{k=0}^{\infty}\rho^{k}\left(\sum_{i=1}^{n}\mu_{i,s-k-1}\right)\right)\right]\end{aligned}\tag{124}$$

Use$\sum_{s=1}^{T}\rho^s = \rho\frac{1-\rho^T}{1-\rho}$ is true is false

$$\begin{gathered} C = \mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\sum_{k=0}^{\infty}\rho^k\alpha_{\tau-k}\right)\right] = \\ \mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t+s-k}\right)\right] = \\ \mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\sum_{k=0}^{s-1}\rho^k\alpha_{t+s-k} + \sum_{k=s}^{\infty}\rho^k\alpha_{t+s-k}\right)\right] = \\ \mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\sum_{k=s}^{\infty}\rho^k\alpha_{t+s-k}\right)\right] = \\ \mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\rho^s\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\right] = \\ \rho^s\sigma_\alpha^2\frac{1}{1-\rho^2}. \end{gathered} \tag{125}$$

XXXXXXXXXXXXXXXXXXXXXXXX

From equation 44

$$\mathbf{E}\left[\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left(\sum_{k=0}^{\infty}\rho^k\alpha_{\tau-k}\right)\right] = \rho^s\sigma_\alpha^2\frac{1}{1-\rho^2} \tag{126}$$

xxxxxxxxxxxxxxxx

$$F_0 = \mathbf{E}\frac{1}{T}\left(\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\left[\sum_t\left(\rho\sum_{k=0}^{\infty}\rho^k\alpha_{t-k}\right)\right] \tag{127}$$

$$F_1 = \mathbf{E}\frac{1}{T}\left(\rho\sum_{k=0}^{\infty}\left(\rho^k\frac{\sum_i\mu_{i,t-k-1}}{n}\right) + \mu_{jt}\right)\left[\sum_t\left(\rho\sum_{k=0}^{\infty}\rho^k\left(\frac{\sum_i\mu_{i,t-k-1}}{n}\right)\right)\right] \tag{128}$$

$$F_2 = \mathbf{E}\frac{1}{T}\left(\rho\sum_{k=0}^{\infty}\left(\rho^k\frac{\sum_i\mu_{i,t-k-1}}{n}\right) + \mu_{jt}\right)\left[\sum_t\left(\mu_{it}\right)\right] \tag{129}$$

C.7 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 stavins96 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 **??**, 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.