Rewarding countries for public spending on clean energy research: Optimal mechanisms under exogenous budgets

Lennart Stern*

February 8, 2022

The latest version can be found here.

Abstract

In 2015, 24 governments created the Mission Innovation Initiative, each pledging to double its public spending for clean energy RD&D between 2015 and 2020. These pledges have only been partially fulfilled. This paper studies how a new global institution could use a given budget to best incentivise countries to expand their clean energy RD&D spending. I develop a static model where each country faces decreasing marginal returns from clean energy RD&D spending and a global institution offers reward payments to countries with the aim to maximize the rate of aggregate research progress. I calibrate the model using the countries' rates of spending on clean energy RD&D observed in 2019. I consider anonymous mechanisms, where the reward payment that a country receives equals its GDP times a function that only depends on its clean energy RD&D spending per GDP. The optimal anonymous mechanism substantially outperforms the Proportional Subsidy Mechanism in which countries are simply reimbursed a fixed fraction of their spending.

1 Introduction

This paper studies how a new global institution with an exogenous budget and the mandate to maximise worldwide progress on clean energy RD&D should best be designed. Further below in this introduction I argue based on related literature that once such an institution exists it might be possible to establish mechanisms generating substantial funding for it. However, first I take this premise as given and summarize the methods and findings of this paper.

A country's spending on clean energy RD&D is an impure global public good (Sandler (2001)): The country reaps private benefits in the form of future

*Paris School of Economics, e-mail: lennart.stern@psemail.eu
For helpful suggestions and comments, I thank Jean-Christophe Poudou.

profits from local companies that the spending might spawn. But the spending can also create benefits to other countries in the form of knowledge spillovers and, arguably most importantly, by reducing global carbon emissions.

Recognizing the important positive global externalities from clean energy RD&D, 24 governments established in 2015 the Mission Innovation Initiative. Members agreed to double their annual clean energy RD&D public expenditures between 2015 and 2020 Myslikova & Gallagher (2020). However, IEA data suggests that MI members increased public investment in clean energy RD&D by \$2.6 billion from 2015 to 2018—far less than the \$4.6 billion they self-reported, and well below the \$9 billion they would need to meet their goal by 2020 Cunliff (2019).

To increase the incentives to spend on clean energy RD&D, a global institution could be created that would reward countries on the basis of their spending. This paper studies how such an institution should best be designed. For this purpose, I develop a simple model in which the rate of progress on clean energy RD&D is a constant-elasticity function of the rate of spending by the country. Based on Acemoglou et al (2016)'s estimate from patent data, I assume an elasticity of $\frac{1}{2}$: Increasing spending by 1% increases progress by 0.5%.

In the model, countries differ in their "absorptive capacity" for clean energy RD&D spending. Countries derive private benefits from progressing on clean energy RD&D but countries with less absorptive capacity run into decreasing marginal returns already at lower levels of spending. The countries' observed spending rates as of 2020 uniquely pin down their types (i.e. the values of their absorptive capacity parameter) in the model. In particular, the model thus generates predictions about how countries would respond to reward payments that the global institution could offer.

The global institution's objective is to maximise the aggregate rate of research progress, summed over all countries. I consider anonymous mechanisms where the global institution cannot taylor its reward payment functions to the individual countries but instead can only condition them on the country's spending per GDP. Specifically, I impose that the reward payment to a country has to equal the product of the country's GDP times a function that only depends on a country's clean energy RD&D spending per GDP.

In this setting, I characterize the optimal anonymous mechanism. I use a version of the model with a continuum of types. For this I provide a sufficient condition for the set of participating types to always be given as an upper interval. This condition is satisfied for the log-uniform type distribution which happens to fit the data quite well. Assuming this distribution, I calibrate the model fully and compute the optimal anonymous mechanism.

I compare this optimal anonymous mechanism to a much simpler alternative: the Proportional Subsidy Mechanism. This alternative mechanism would simply reimburse all countries a fixed fraction of their spending on clean energy RD&D. At an annual budget of \$2billion, the Proportional Subsidy Mechanism achieves 63.5% of the welfare gains that the optimal anonymous mechanism achieves and at \$1billion only 53.0%.

How valuable would it be to create such a fund rewarding countries for

public spending on clean energy RD&D? This paper gives an answer to this question conditional on any given level of funding at the global fund's disposal. For example, for an annual budget of 2 billion, the optimal anonymous mechanism causes aggregate spending on clean energy RD&D to increase by 3.1 times that amount, leading to an acceleration of the global aggregate rate of research progress to increase by 26%.

This formal analysis is incomplete, since I will not provide any model with endogenous funding. However, I will now argue based on related literature that the creation of such an instution might open up the possiblity for establishing mechanisms to generate substantial funding for it.

A sceptical view about the value of creating a fund for rewarding countries for public spending on clean energy RD&D could be expressed as follows: Why would such a Global Public Good Institution (GPGI) receive any funding? Currently, global welfare could be increased if all countries increased their spending on clean energy RD&D. Each country fails to internalise the benefits of this research on other countries and thus chooses spending levels that are suboptimally low from the perspective of global welfare. Given that thus countries manifestly fail to cooperate in this game, why would they manage to cooperate in the game of contributing to the fund for rewarding countries for public spending on clean energy RD&D?

I will now provide 3 responses to this sceptical view. Firstly, the global institution proposed in this paper could have a high marginal cost-effectives at least for small budgets. This is because since at the status quo, each country's level of spending on clean energy RD&D corresponds to an interior optimum of its maximisation problem. Thus its marginal loss from deviating a bit from its optimum is of second order. This implies that the marginal cost-effectiveness of the money reaching the institution is infinite when its budget is 0. As a result, countries will have incentives to contribute at least some money to the global institution.

Secondly, the global institution proposed in this paper could receive funding in an analogous way to the existing Global Public Good Institutions (GPGIs). In fact, the 14 billion of funding that GPGIs are currently receiving per year (see Stern (2020)) consist almost entirely of voluntary contributions that are announced and reviewed by countries at dedicated replenishment conferences. If a new GPGI like the one proposed in this paper was created, such public pledging rounds for funding could be held at the recurring Mission Innovation review conferences.[1]

[1] These replenishment conferences provide countries effective opportunities to signal altruism. In fact, when a country spends on clean energy RD&D, it is never entirely clear to what extent this is for altruistic reasons or for self-interested reasons based on the local benefits that the country internalises from research happening on its terrritory. When it comes to contributing to a GPGI, it is much clearer that the contributions are at least partly altruistically motivated, at least when the overall funding level are large. In Stern (2020) I have argued that the current pattern of voluntary contributions to GPGIs, where each GPGI has at least 4 and typically more than 10 substantial donors, is inconsistent with a model where countries have purely consequentialist objective functions. It is, however, consistent with the informal "signalling of altruism model" that I sketched just now.

A third and arguably most important consideration weighing in favor of creating new GPGIs like the one proposed in this paper is that this would open up the possiblity for establishing new mechanisms for funding GPGIs.

For example, Molina et al. (2020) study a two stage matching mechanism where in the first stage each country i announces for each other country j a "matching factor" m_{ij} and in the second stage an unconditional contribution a_i. The rules of the scheme stipulate that each country pays in the end the amount $a_i + \sum_j m_{ij} a_j$. Molina et al. (2020) show for the case of 2 players that the subgame perfect Nash equilibrium in the induced game leads to globally efficient contributions.

Arguably, this mechanism can be applied well to monetary contributions to a global institution like the one proposed in this paper. The first stage (i.e. the announcement of the matching factors m_{ij}) could be held at the beginning of each year and the voluntary contributions made by country during this year could be regarded as the unconditional contributions (i.e. the a_i), so that on their basis the matching contributions, $\sum_j m_{ij} a_j$, would be determined at the end of the year.

This mechanism could of course also be applied to countries spending on clean energy RD&D (instead of applying it to their contributions to the global institution proposed in this paper that would reward countries on the basis of their spending). However, this might result in drastic fluctuations in countries' spending which could be quite costly. On the other hand, fluctuations in the funding available to the global institution proposed in this paper would result in less sudden fluctuations in countries' spending. Thus, the mechanism from Molina et al. (2020) might be better applied to countries' monetary contributions to the global institution proposed in this paper rather than directly to countries' spending choices for clean energy RD&D. As a result, the creation of the global institution proposed in this paper could be useful in that it could facilitate the solution of the countries' collective action problem.

Another example[2] of how the creation of global compensation funds like the one proposed in this paper could prepare the ground for establishing mechanisms to increase the provision of global public goods is provided by Kornek and Edenhofer (2020). They define a class of club mechanisms where membership requires certain contributions to a compensation fund. Only members are eligible to receive reward payments from the compensation fund. Kornek and Edenhofer (2020) show for a certain version of the compensation fund that if membership dues are made proportional to the size of the set of participants then the subgame perfect Nash equilibrium of the the induced game achieves global welfare maximisation.

Furthermore, simple mechanism could be devised that strategically harness the terms of trade effects of different forms of climate change mitigation to induce countries to contribute. For this it would be important to create two separate funds of the type studied in this paper: Firstly, a fund rewarding

[2] Another category of possible mechanisms are those based on global resources like the IMF's special drawing rights (Aryeetey (2004), Stiglitz (2006), Birdsall and Leo (2011)) or the revenue from deep seabed mining beyond national jursidiction Feichtner (2019).

countries on the basis of their per GDP spending on all clean energy research that in expectation increases future demand for fossil fuel. This would include all RD&D on Carbon Capture and Storage (CCS), since cheaper CCS will mean that carbon taxes will lead more to CCS expansion instead of reduction in fossil fuel use.

Secondly, a separate fund should be created that would reward countries on the basis of their per GDP spending on all clean energy research that in expectation decreases future demand for fossil fuel. This includes all research on renewable energy.

A Proportional Matching Fund (PMF) could then be created that would at the beginning of each year collect donations. At the end of each year, the PMF would split its budget between the two funds described above in proportion to aggregate direct donations made to them during the year.

Fossil fuel exporters would then have an incentive to donate money to the fund rewarding countries for the research that increases future demand for fossil fuels, since this will increase the price of fossil fuels. Conversely, fossil fuel exporters would then have an incentive to donate money to the fund rewarding countries for the research that increases future demand for fossil fuels, since this will increase the price of fossil fuels.

These incentives to donate are enhanced through the Proportional Matching Fund (PMF). As a result, countries primarily motivated by reducing climate change have an incentive to give to the PMF, since by doing so they can induce fossil fuel exporters and importers to give more to their preferred funds.

Assessing the potential of such a scheme would require making assumptions about how current research progress translates into future reductions in fossil fuel demand which is beyond the scope of this paper. However, related work (Stern (2022)) gives reason for optimism: In the case of a PMF for funds that reward countries for taxing the production and the combustion of coal, respectively, $11.84 billion get raised at the Subgame Perfect Nash Equilibrium.

The current paper tries to complement this literature by providing a fully operational proposal for the design of a compensation fund to induce countries to increase their spending on clean energy RD&D. The above-mentioned papers give reasons to hope that once new scalable compensation funds like the one proposed in this paper are established it might be possible to devise new mechanisms to generate substantial funding for them.

2 Related Literature

This paper builds on the theory of incentives (Laffont and Martimort (2009)) by applying a simple model with asymmetric information to the question of how to best incentivise countries to increase their spending on clean energy RD&D. I follow Martimort and Sand-Zantman (2013,2016) in studying anonymous mechanisms for a global institution contracting with countries and in using a continuum of types. I provide a detailed discussion of the advantages of anonymous mechanisms in section 8. Concerning the use of a continuum of

types, the motivation is as follows: The use of a discrete type model fitted to the observed data presented in section 4.2 would bring with it the risk of a kind of "overfitting". The optimal mechanism would be finely taylored to the currently observed data and rely sensitively on the accuracy of the model.

To avoid this sensitivity, I assume a simple functional form (log-uniform, see section 4.2) for the distribution of the types. The functional form is chosen so as to provide a good approximation to the distribution of types inferred from the data for countries spending on clean energy RD&D in 2019 (see section 4.2). This approach results in a smooth optimal reward function. In contrast to this, a discrete type model would only yield an implementation with a discrete menu of choices and associated reward payments that countries could choose from. Given that the model will unlikely to be very accurate, it is likely to be better in practice to use continuous reward payment functions like the ones found to be optimal in section 7.

3 The model

Each country is characterized by its GPD, denoted y, and a coefficient, denoted a, that captures its research institutions' capacity to absorb spending on clean energy research. This reflects the quantity and quality of people with relevant knowledge and skills.

Each country also has a parameter β that gives how much it values a unit of progress on clean energy RD&D. However, I will assume that this parameter is the same for all countries. This is motivated by an assumption that countries value such progress primarily because of the private benefit in the form of profits reaped by spinoff companies that commercialise the research findings. A more complete model would account for the fact that countries also value the climate change mitigation benefits of clean energy RD&D, either due to the partial internalisation of the climate change mitigation benefits (for large countries) or due to partial altruism. However, this is beyond the scope of the current paper.

The model is static and each country's only choice variable is the expenditure x on clean energy research. The country's utility is given by u as follows:

$$u = \beta a(\frac{x}{y})^{\lambda} y - x$$

The term $a(\frac{x}{y})^{\lambda} y$ is meant to correspond to the rate of progress on clean energy research in the country. I assume that $\lambda \in (0, 1)$. The specification means that increasing the rate of spending by 1% increases the rate of progress on this research by λ%. I thus assume that this elasticity is constant and the same for all countries. The parameter a can be interpreted to be the "absorbtive capacity" for clean energy research spending in the country relative to its GDP.

For the empirical calibrations I will consider mechanisms where a country's reward payment is given as $yt(\frac{x}{y})$, where t is a function that only depends on countries' research spending relative to GDP. As a result, a country's payoff will be:

$$\beta a(\frac{x}{y})^{\lambda}y - x + yt(\frac{x}{y}) = (\beta a(\frac{x}{y})^{\lambda} - \frac{x}{y} + t(\frac{x}{y}))y$$

Thus a country's problem boils down to choosing the fraction $\frac{x}{y}$ of GDP to spend on clean energy research and the optimal such choice only depends on βa. This observation will allow us to study a simplified version of the model where y is the same for all countries and then later apply it to the actual empirical data. For ease of exposition, I will from now on consider the model where there is a continuum of countries of measure 1 with all countries having the same GDP, also normalized to 1.

A country's utility is thus given as follows:

$$u(x, \beta a) = \beta a x^{\lambda} - x$$

Since I am assuming that β and λ are the same across all countries, we could thus view the parameter a as being the country's type. However, for the algebra it will be convienent to instead view the type as being the following positive and strictly increasing transformation of a:

Definition 1.

$$b := (\beta a \lambda)^{\frac{1}{1-\lambda}}$$

Instead of reasoning about the absolute utility of a type b, it will be convenient to reason about its utility relative to not participating in the mechanism. We denote this by v:

Definition 2.

$$v(x, b) := u(x, \frac{b^{1-\lambda}}{\lambda}) - sup_{x'} u(x', \frac{b^{1-\lambda}}{\lambda})$$

Since the country's utility is quasilinear in money, its overall payoff relative to not participating in the mechanism is

$$v(x, b) + t$$

where t denotes the transfer that the country receives.

I assume that there is a continuum of types with cumulative distribution function $F(b)$ with support given by the interval $[\underline{b}, \bar{b}]$. The corresponding probability density function is denoted $f(b)$.

By definition 1, we have:

$$u(x, \frac{b^{1-\lambda}}{\lambda}) = \frac{b}{\lambda}(\frac{x}{b})^{\lambda} - x$$

From the definition 2, we directly deduce:

Lemma 1. *We have:*

$$v(x, b) := \frac{b}{\lambda}((\frac{x}{b})^{\lambda} - 1) - (x - b)$$

$$\frac{\partial v}{\partial b} = (x^\lambda - b^\lambda)\frac{1-\lambda}{\lambda b^\lambda}$$

Moreover, $argmax_{x'} v(x', b) = b$ is the type b's spending on clean energy RD&D in the absence of any mechanism.

Now I introduce the global institution. This global institution has an exogenous budget M. Its objective is to maximise the aggregate rate of progress on clean energy research:

$$w = \int_{b=\underline{b}}^{\bar{b}} a x^\lambda f(b) db$$

Now let us multiply both sides by β. Recall that β denotes the utility that countries derive per unit of domestic progress on clean energy RD&D and that I am assuming that this has the same value for all countries. Maximizing w is the same problem as maximizing βw:

$$\beta w = \int_{b=\underline{b}}^{\bar{b}} \beta a x^\lambda f(b) db$$

This we can write in terms of the type b:

$$\beta w = \int_{b=\underline{b}}^{\bar{b}} \frac{b^{1-\lambda}}{\lambda} x^\lambda f(b) db$$

From now on, I will take this to be the global institution's objective. The rest of this section motivates the assumption that the global institution only cares about increasing the aggregate rate of progress on clean energy RD&D.

3.1 Motivation of the assumed objective of the global institution

I will now informally motivate this specification by arguing that it is a good approximation to a more complete model where the global institution cares about aggregate global welfare in a time-neutral way. In such a more complete mode, the global institution would thus take into account both long-term welfare and short term welfare. It affects long-term welfare mostly via the reduction in climate change damages resulting from the acceleration of clean energy RD&D. Short term welfare includes the countries' utilities, $v(x, b)$. Thus our assumption that global welfare w simply equals the acceleration in clean energy RD&D roughly amounts to neglecting the "short term utilities", $v(x, b)$. However, this approximation is unlikely to greatly affect the optimal mechanism, as I will now argue.

Even without taking the utilities $v(x, b)$ into account in the global institution's objective, they do enter its Lagrangian via the multiplier associated with the budget constraint: The lower the utilities $v(x, b)$, the larger the required transfers that the global institution needs to pay countries to make them

participate. Thus also taking into account the utilities $v(x, b)$ in the global institution's objective merely increases the relative weight given to them in the Lagrangian. If the Lagrange multiplier associated with the budget constraint is large then this will not change the Lagrangian by much.

Thus we have informally established that the approximation used in the current paper is good as long as the Lagrange multiplier associated with the budget constraint is large. But is it plausible that this Lagrange multiplier is indeed large? In other words, should we expect the marginal value of money going to a global institution rewarding countries for spending on clean energy RD&D to be large? Arguably, the answer is yes, given that countries are unlikely to manage to reach a fully cooperative outcome and given their short-termist bias.

This conclusion is in fact born out in a related paper (Stern (2021)) that studies how to best reward countries for taxing fossil fuels. There, I find that under a social cost of carbon of \$417/tCO2 the mechanism that maximizes global welfare is very well approximated by the mechanism that simply maximizes global emission reductions, as long as the global institution's annual budget does not greatly exceed \$100 billion.

By analogy, it is plausible that in the current paper the mechanism that maximizes global (time neutral) welfare is very well approximated by the mechanism that maximizes the aggregate progress of clean energy RD&D. It is easy to see that this claim must hold for sufficiently small budgets. To check that it holds for realistic values of the annual budget of the global institution proposed in this paper (e.g. for <\$10 billion per year), it would be necessary to make assumptions about the impact of clean energy RD&D on long term cumulative carbon emissions. I leave this task for future work.

4 Empirical Calibration

4.1 The elasticity of research spending

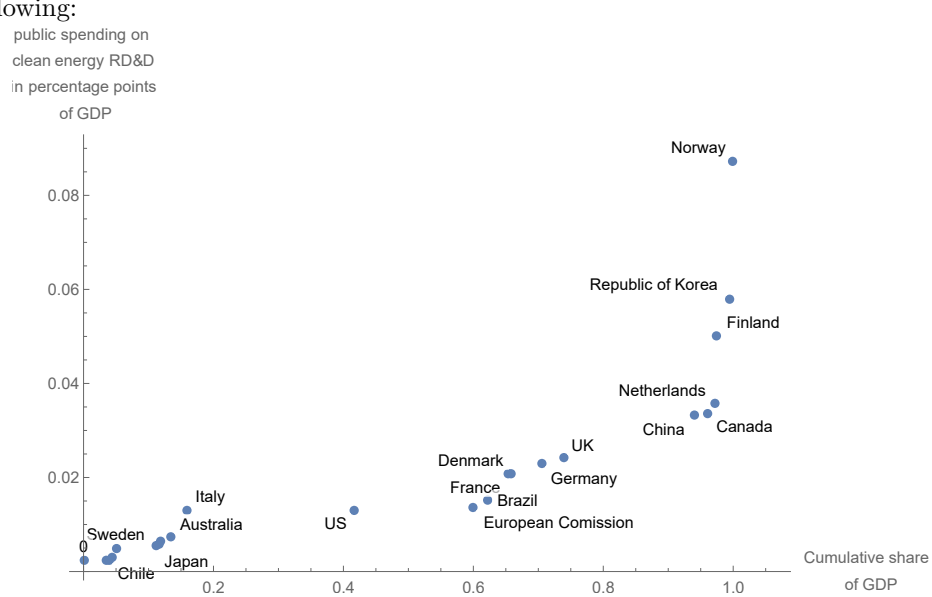
The parameter λ tells us by what percentage research is accelerated if the funding flow is increased by 1%. Using patent data, Acemoglou et al (2016) estimate this to be $\frac{1}{2}$. This means that doubling the rate of research progress requires quadrupling the rate of funding. I will later specialize the model to this case:

Assumption 1. *The funding elasticity of research progress λ is $\frac{1}{2}$.*

4.2 The status quo distribution of research spending

By Lemma 1, the type b equals the corresponding country's spending on clean energy research in the absence of any mechanism. In the model as written above, all countries are of equal size (in the sense of GDP). However, for applying the model to the actual world with heterogenous country sizes, we need to reinterpret b to equal the spending relative to GDP, as explained above.

The cumulative distribution function $F(b)$ gives the mass of countries whose spending on clean energy RD&D per GDP in the absence of any mechanism is less than b. The countries participating in Mission Innovation report this data in a standardized way on an annual basis. For the other countries, I have not found this data and for this reason I have restricted the sample of countries to the participants in Mission Innovation. In 2020, these countries[3] spent the following:

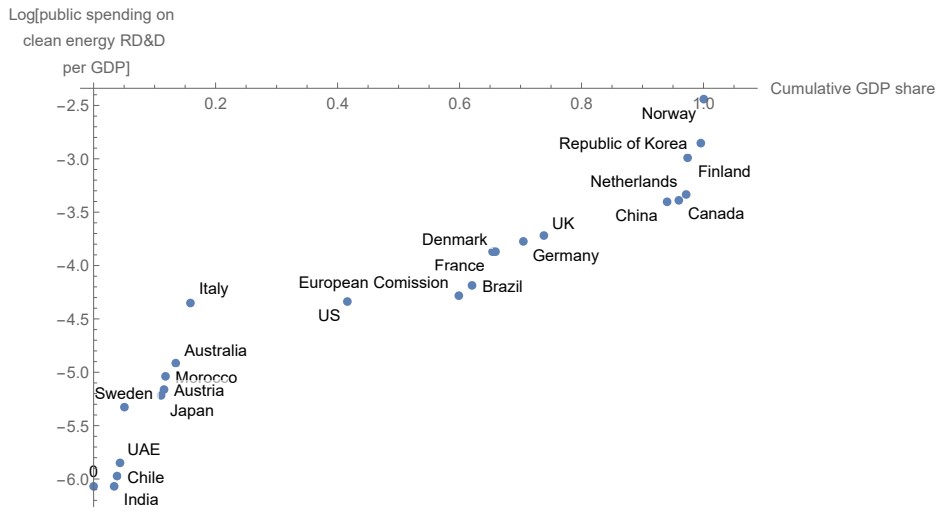


This plot is obtained by first writing down the countries in ascending order of their per GDP spending on clean energy RD&D. The plotted point corresponding to a country i has as x-coordinate the share of GDP comprised by countries spending at most as much as country i per GDP on clean energy RD&D. The y-coordinate is country i's spending on clean energy RD&D per GDP.

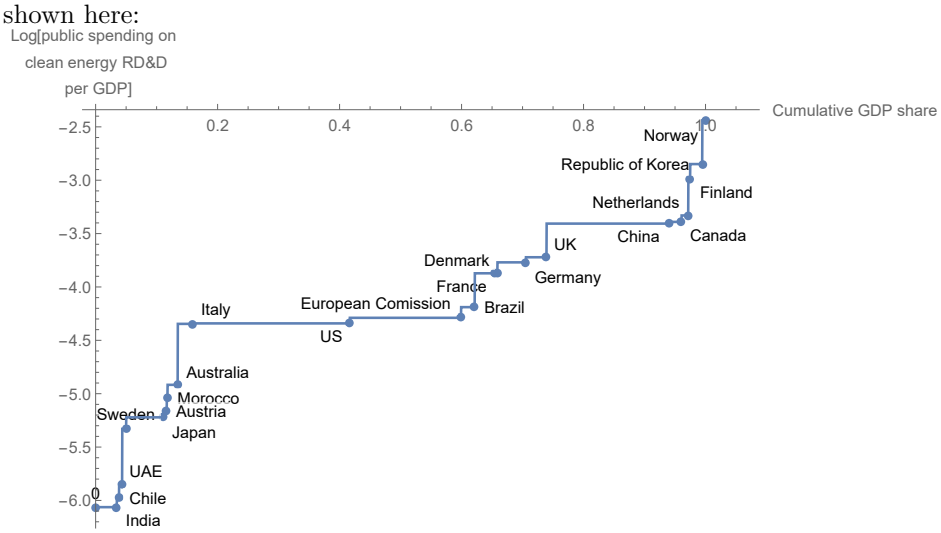
So far, there does not exist any mechanism rewarding countries for spending on clean energy RD&D. Thus the plotted values correspond to the values of b by Lemma 1. As a result, we can practically read off the empirical cumulative distribution function $F(b)$ from the plot. We want to approximate this distribution function using a simple functional form. From the above plot, it remains unclear which functional form is appropriate.

However, plotting instead the natural logarithm of the spending per GDP yields the following:

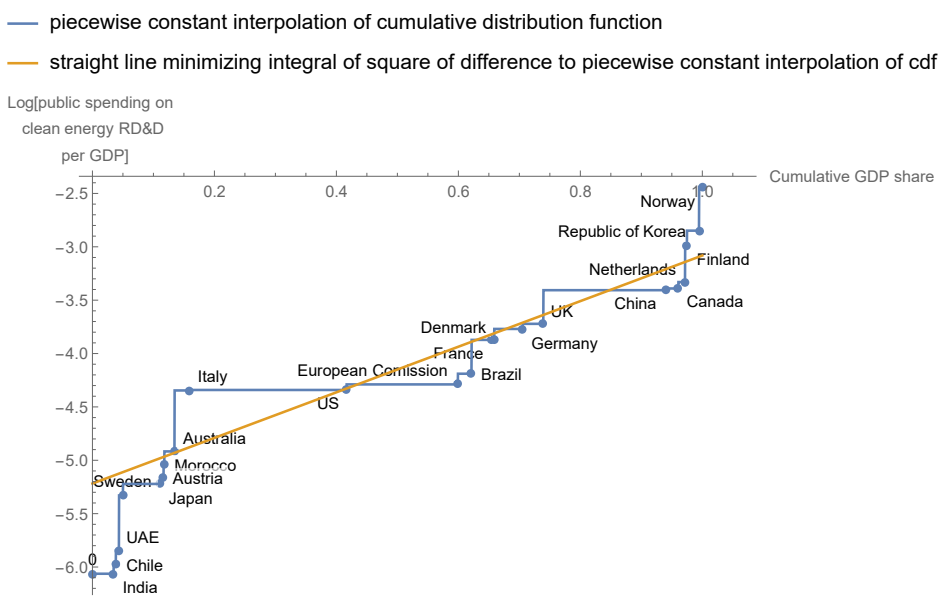
[3]The European Comission participates in Mission Innovation as an actor in its own right. It would be natural to allow for this in the definition of the reward payment schemes and simply treat such actors as if they were a country with GDP equal to the sum of the GPDs of the countries that they comprise. This is how I am treating the European Comission here for the empirical calibration of the model.



From this we can construct (the inverse of) a cumulative distribution function as shown here:



For the purpose of applying the model, I will use a continuum of types. The straight line shown in the following plot is the straight line that minimises the integral of the square of the difference to the function obtained from the data:



The graph suggests that this gives a fairly good approximation to the empirically observed distribution. This motivates the following assumption:

Assumption 2. $F(b) = \frac{log(b)-log(\underline{b})}{log(\bar{b})-log(\underline{b})}$, *where* $(\underline{b}, \bar{b})$ *is chosen so as to minimize the integral of the square of the difference to the empirically generated distribution function.*

4.3 The status quo aggregate spending on clean energy RD&D

In 2020, the aggregate public spending on clean energy RD&D by the participants in Mission Innovation was \$16.06 billion. In the model, this corresponds to the spending in the absence of any global institution, which is $\int_{b=b_L}^{b_H} F'(b)bdb$. I use this to complete the calibration of the model.

5 The optimal mechanism under complete information

Under complete information, the global institution's problem is to choose a spending level $x(b)$ for each type b to solve the following:

Problem 1.

$$\max_{x(.)} \int_{b=\underline{b}}^{\bar{b}} \frac{b^{1-\lambda}}{\lambda}(x(b))^{\lambda} f(b)db$$

under the following 2 constraints:

1. exogenous budget:

$$\int_{b=\underline{b}}^{\bar{b}} t(x(b))f(b)db \leq M$$

where M denotes the global institution's exogenous budget.
2. voluntary participation:

$$v(x(b), b) + t(x(b)) \geq 0$$

The optimal mechanism under complete information turns out to scale up research spending in all countries by the same factor, as shown in the following Lemma:

Lemma 2. *Under complete information, the optimal contract for a given budget M at the global institution's disposal implements the following profile of spending rates:*

$$x(b) = \frac{b\left(\sqrt{M\log\left(\frac{\bar{b}}{\underline{b}}\right)} + \sqrt{\bar{b}-\underline{b}}\right)^2}{\bar{b}-\underline{b}}$$

The resulting aggregate progress on clean energy RD&D is given by:

$$2(\bar{b}-\underline{b})\frac{\left(1+\sqrt{\frac{M\log\left(\frac{\bar{b}}{\underline{b}}\right)}{\bar{b}-\underline{b}}}\right)^2}{\log\left(\frac{\bar{b}}{\underline{b}}\right)}$$

Proof. See the accompanying Mathematica file. □

Consider $\frac{d(x(b))}{dM}$, i.e. the marginal increase in research spending induced in type b per additional dollar at the global institution's disposal. We compute:

$$\frac{dx(b)}{dM} = \frac{\sqrt{\log\left(\frac{\bar{b}}{\underline{b}}\right)}}{\sqrt{M}}\frac{b\left(\sqrt{M\log\left(\frac{\bar{b}}{\underline{b}}\right)} + \sqrt{\bar{b}-\underline{b}}\right)}{\bar{b}-\underline{b}}$$

For small M, the marginal increase in research spending induced in type b, namely $\frac{d(x(b))}{dM}$, behaves like $\frac{1}{\sqrt{M}}$. In particular, $lim_{M\to 0}\frac{d(x(b))}{dM} = \infty$. Put differently: The marginal cost of inducing the type b to increase its spending by an infinitesimal amount is 0 when $M = 0$. This result is true much more generally: If the agents' optimization problem in the absence of any mechanism has an interior solution then their cost of deviating slightly is of second order.

Now let us compute the "marginal crowding-in factor", defined as the aggregate additional spending induced by an additional dollar at the global institution's disposal. Recalling that $F(b) = \frac{log(b)-log(\underline{b})}{log(\bar{b})-log(\underline{b})}$ and thus $f(b) = \frac{\frac{1}{b}}{log(\bar{b})-log(\underline{b})}$, we obtain:

$$\int_{b=\underline{b}}^{\bar{b}} f(b)\frac{dx(b)}{dM} = 1 + \frac{1}{\sqrt{M}}\sqrt{\frac{\bar{b}-\underline{b}}{\sqrt{log(\bar{b})-log(\underline{b})}}}$$

In particular, this is decreasing in M and we also get:

$$\int_{b=\underline{b}}^{\bar{b}} f(b)\frac{dx(b)}{dM} > 1$$

$$lim_{M\to\infty}\int_{b=\underline{b}}^{\bar{b}} f(b)\frac{dx(b)}{dM} = 1$$

These properties simply follow from the fact that countries derive positive and decreasing marginal benefits from domestic clean energy RD&D, with marginal benefits converging to 0 as the spending goes to infinity.

6 The optimal mechanism under incomplete information

In this section, I will study the optimal mechanism under incomplete information, by which I mean here the optimal mechanism for when the global institution knows the distribution of types but can condition its reward payments only on countries' actions, i.e. on their rates of spending on clean energy RD&D. I will also synonomously refer to the optimal mechanism in this setting as "the optimal anonymous mechanism".

The motivation for this is as follows. In practice, the countries' types are their private information. These types can change over time, so a country's past choices do not fully reveal its type. Moreover, there are important advantages from only conditioning reward payments on the countries' actions which I discuss in detail in section 8.

6.1 The global institution's problem under incomplete information

Under incomplete information (or in the setting we are interested in where the global institution can only condition on the spending rate x but not the type b), the following incentive compatibility constraint arises:

$$v(x(b), b) + t(x(b)) \geq v(x', b) + t(x')\forall b, x'$$

This is because the type b is free to choose any x, so its choice $x(b)$ cannot be worse for it than any other x.

Definition 3. Let us define $U(b)$ as the utility gain that type b can realise by participating relative to not participating (in which case its optimal research spending rate is $x = b$.). We thus have:

$$U(b) := sup_x v(x,b) + t(x)$$

By the envelope theorem, we have

$$U'(b) = \frac{\partial v}{\partial b}|_{(x(b),b)}$$

where $x(b) := argmax_x\, v(x,b) + t(x)$.
Using Lemma 1, we get:

$$U'(b) = (x(b)^\lambda - b^\lambda)\frac{1-\lambda}{\lambda b^\lambda}$$

I will show in the proof of Lemma 5 that it is never optimal for the global institution to implement an $x(b)$ with $x(b) < b$ for any b. Thus we have that $U(b)$ is non-decreasing. Hence there exists b_{min} such that all $b < b_{min}$ do not participate whilst all $b > b_{min}$ do participate. Moreover, the participation constraint must be binding at b_{min}. Thus we obtain:

$$U(b) = 0 \forall b \leq b_{min}$$

$$U(b) = \frac{1-\lambda}{\lambda}\int_{z=b_{min}}^{b} (\frac{x(z)^\lambda - z^\lambda}{z^\lambda})dz \forall b \geq b_{min}$$

We also have:

$$U(b) = v(x(b),b) + t(b)$$

which yields:

$$t(b) = -v(x(b),b) + U(b)$$

Lemma 3. *Any implementable $x(b)$ is non-decreasing.*

Proof. See appendix A.1. □

Thus the global institution's problem reduces to the following:

Problem 2.

$$\max_{x(.)} \int_{b=\underline{b}}^{\bar{b}} f(b)\frac{b^{1-\lambda}}{\lambda}x(b)^\lambda db$$

under the following constraint:

$$\int_{b=\underline{b}}^{\bar{b}} f(b)(-v(x(b),b) + U(b))db \leq M$$

where M denotes the global institution's exogenous budget and $U(b)$ is given by:

$$U(b) = 0 \forall b \leq b_{min}$$

$$U(b) = \frac{1-\lambda}{\lambda} \int_{z=b_{min}}^{b} \frac{x^{\lambda} - z^{\lambda}}{z^{\lambda}} dz \forall b \geq b_{min}$$

Lemma 4. *Suppose that $\frac{bf'(b)}{f(b)}$ is non-increasing. Then at the optimal mechanism under a given budget and incomplete information there exists b_{min} such that all $b < b_{min}$ do not participate whilst all $b > b_{min}$ do participate. In particular, this holds when b is distributed log-uniformly.*

Proof. See appendix A.2. □

Let us now interpret the meaning of the condition that $\frac{bf'(b)}{f(b)}$ be non-decreasing. This can be viewed as an elasticity: It gives by what percentage the density changes if we increase the type b by 1%.

Lemma 5. *Suppose assumption 1 holds, i.e. suppose that the elasticity of research spending λ is $\frac{1}{2}$. For a given budget M at the global institution's disposal, the second best optimal allocation $\tilde{x}(b)$ is as detailed below.*

Define $\tilde{M} := \frac{b}{4}\left(\frac{\bar{b}}{\underline{b}}\log(\frac{\bar{b}}{\underline{b}}) - \frac{2(\frac{\bar{b}}{\underline{b}}-1)}{\log(\frac{\bar{b}}{\underline{b}})} + 2\right)$. For $M < \tilde{M}$ the second best optimal mechanism involves participation only by the types above a value b_{min} defined below, whilst for $M \geq \tilde{M}$, the optimal mechanism involves full participation.

The Lagrange multiplier $\mu(M)$ associated with the global institution's budget constraint is given by the following function:

For $M < \tilde{M}$, $\mu(M)$ is the unique solution to the following equation:

$$\frac{\bar{b}\left((1+\frac{2}{\mu})e^{-2/\mu} - \left(1 - \frac{2}{\mu^2}\right)\right)}{2\log\left(\frac{\bar{b}}{\underline{b}}\right)} = M$$

For $M \geq \tilde{M}$, $\mu(M)$ is the following expression:

$$\frac{2}{\sqrt{\frac{\log\left(\frac{\bar{b}}{\underline{b}}\right)(4m - \underline{b}(2+\log(\frac{\bar{b}}{\underline{b}})))}{\bar{b}-\underline{b}} + 2}}$$

The second best optimal allocation is given by $\tilde{x}(b)$ defined as follows:

$$\tilde{x}(b) = max(1, (\frac{1}{2}log(\frac{b}{\bar{b}}) + 1 + \frac{1}{\mu(M)})^2)b$$

In particular, we have that $b_{min} = max(\underline{b}, e^{-\frac{2}{\mu}}\bar{b})$, meaning that the types with $b \geq b_{min}$ participate whilst the types with $b \leq b_{min}$ choose $x(b) = b$ just as in the absence of the mechanism. If $b_{min} < \underline{b}$ then all types participate.

This allocation can be implemented through the reward payment function $t(x)$ defined through the following expression:

$$\max\left(0, \frac{b_{min}}{2}\left(\log\left(\frac{\bar{b}}{b_{min}}\right)+1-\frac{2}{\mu(M)}\right)+\frac{b}{2}\left(\log\left(\frac{b(x)}{\bar{b}}\right)\left(\frac{1}{2}\log\left(\frac{b(x)}{\bar{b}}\right)+1+\frac{2}{\mu(M)}\right)-1+\frac{2}{\mu(M)}+\frac{2}{\mu(M)^2}\right)\right)$$

which can also be written as:

$$\max\left(0, (b_{min}-b)(\frac{1}{2}-\frac{1}{\mu(M)})+\frac{b_{min}}{2}\log\left(\frac{\bar{b}}{b_{min}}\right)+\frac{b}{(\mu(M))^2}+b(\frac{1}{2}+\frac{1}{\mu(M)})\log\left(\frac{b(x)}{\bar{b}}\right)+\frac{b}{4}(\log\left(\frac{b(x)}{\bar{b}}\right))^2\right)$$

where $b_{min} := \max\left(\underline{b}, \bar{b}e^{-\frac{2}{\mu(M)}}\right)$ as defined above and where $b(x)$ is defined to be the unique solution for b to the following equation:

$$b\left(\frac{1}{2}\log\left(\frac{b}{\bar{b}}\right)+1+\frac{1}{\mu(M)}\right)^2 = x$$

Proof. See appendix A.2. □

Reassuringly, we can immediately deduce from Lemma 5 that $lim_{\mu\to\infty}\tilde{x} = b$: In the limit as the global instition's budget goes to 0, all types end up choosing the same spending rates as in in the absence of any reward payments.

For the case where b is distributed log-uniformly, Lemma 5 shows that the types that end up participating at the optimal mechanism multiply their spending relative to the situation without any mechanism by a factor that increases in their type.

Lemma 6. *The second best optimal allocation $\tilde{x}(b)$ can be implemented via the reward payment function $t(x)$ defined to be as follows:*

$t(x) = 0$ *for* $x < \tilde{x}(b_{min})$

$t(x) = \int_{z=b_{min}}^{\tilde{x}^{-1}(x)} \frac{\partial v}{\partial b}(\tilde{x}(z), z)dz - v(x, \tilde{x}^{-1}(x))$ *for* $x \in Image(\tilde{x}|_{[b_{min},\bar{b}]})$

$t(x) = \int_{z=b_{min}}^{\bar{b}} \frac{\partial v}{\partial b}(\tilde{x}(z), z)dz - v(\tilde{x}(\bar{b}), \bar{b})$ *for* $x > \tilde{x}(\bar{b})$

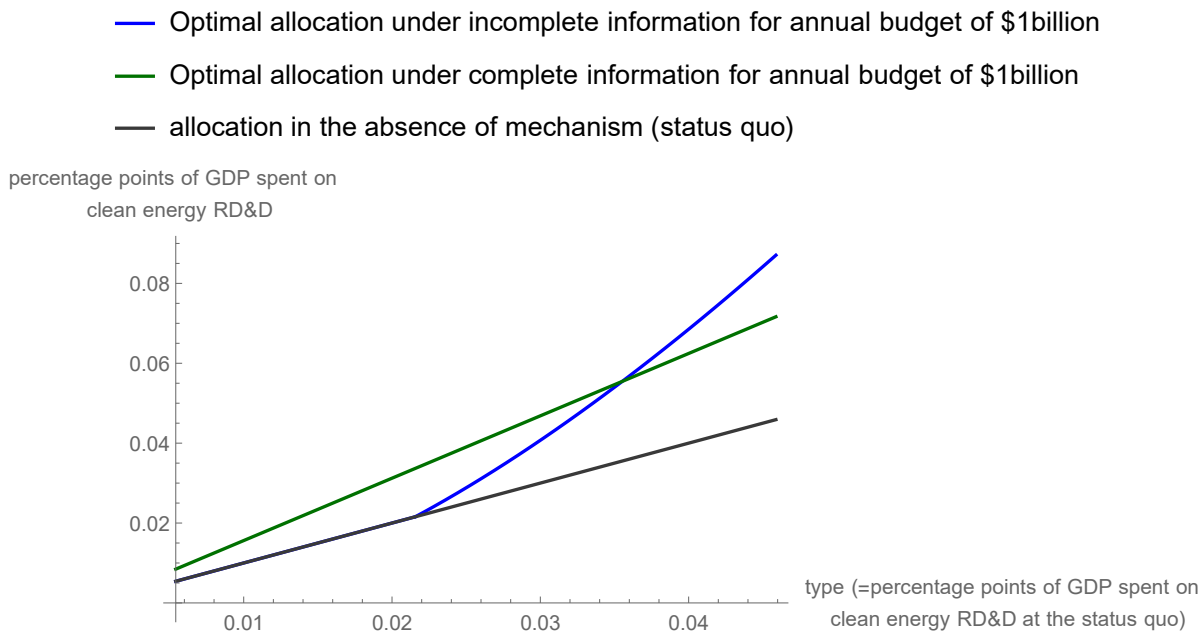
Proof. See appendix A.4. □

7 Numerical Results

I now provide numerical results based on the calibration detailed in section 4.

7.1 The optimal allocation under incomplete information

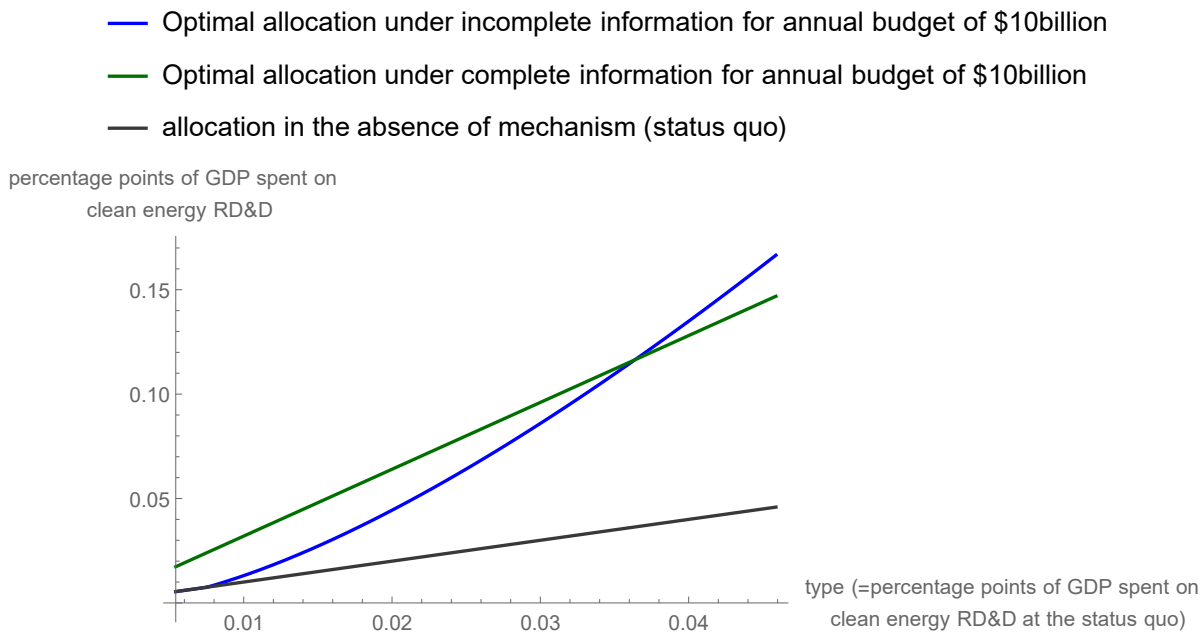
Here is a direct comparison between the optimal mechanism under incomplete and complete information for when the global institution's annual budget is \$1billion:



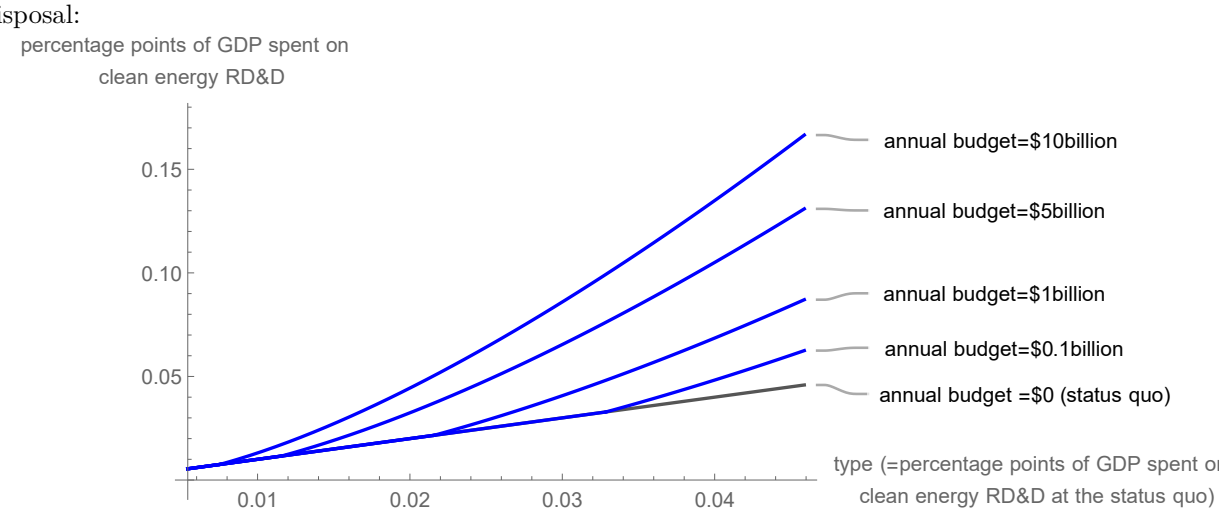
The optimal mechanism under complete information induces all types to scale up their spending on clean energy RD&D by a fixed factor. This is because under complete information the only relevant consideration is efficiency. Uneven scaling up of the spending on clean energy RD&D leads to losses in efficiency due to the decreasing marginal returns that those countries run into that scale up more than the others.

However, under incomplete information there is a further consideration: the reduction in informational rents. This weighs in favor of scaling up more the spending of the higher types because incentivising any given type to do so creates rents for all the higher types. This explains the strict concavity of the optimal allocation under incomplete information.

Here is the case where the global institution's annual budgets is $10billion:



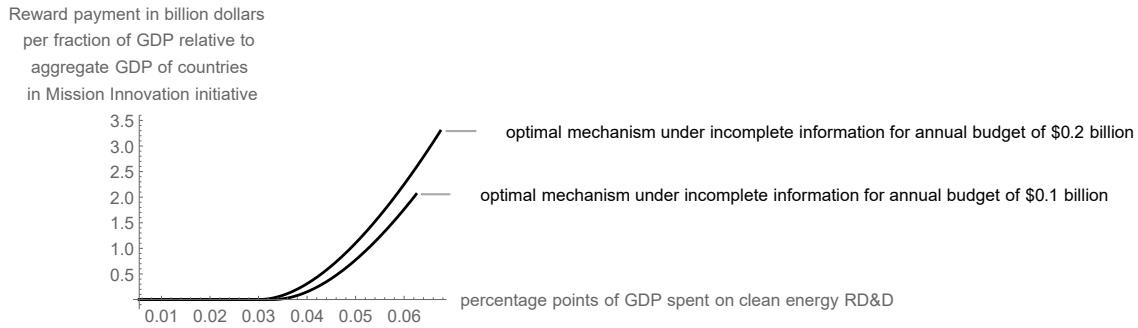
The following graph shows the allocation at the optimal mechanism under incomplete information for different annual budgets at the global institution's disposal:



Recall that by Lemma 1 the type b spends the amount b on clean energy RD&D in the absence of any mechanism. Thus the purple line the above graph is the identity function.

7.2 The reward payment functions implementing the optimal allocation under incomplete information

The global institution can implement the optimal mechanism under incomplete information using the following reward payment functions:

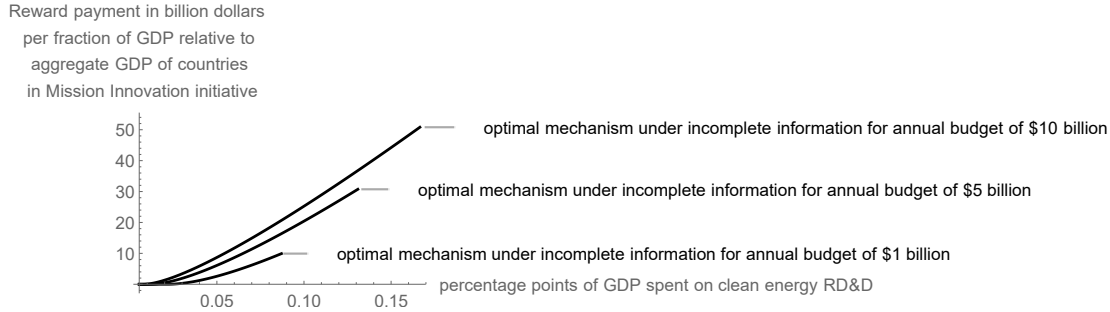


I have plotted on the x-axis the interval starting from the percentage of GDP spent on clean energy RD&D by the lowest type and ending at the percentage of GDP spent by the highest type under the optimal mechanism. Thus the curve for an annual budget of \$0.1billion is plotted up to 0.068, since under the optimal mechanism for this annual budget the highest type ends up spending 0.068% of its GDP.

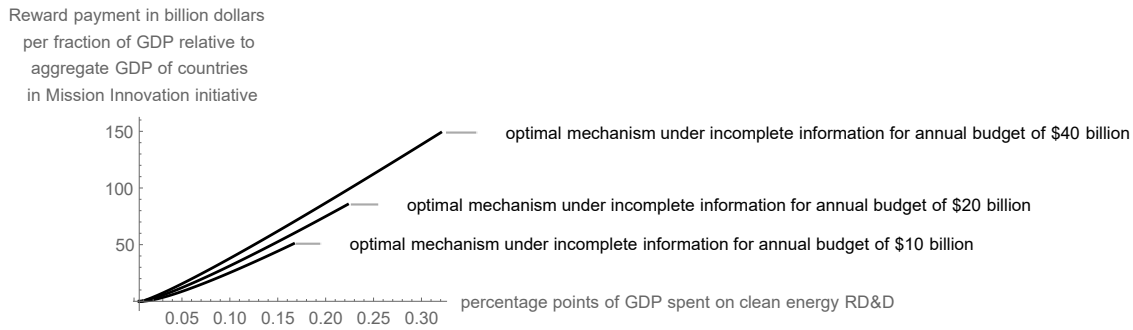
The y-axis is to be interpreted as follows: A country with a GDP making up a proportion q of the aggregate GDP of all the countries receives a reward payment obtained by multiplying the value read off from the above graph by q.

For example, consider the reward payment function corresponding to the optimal mechanism under an annual budget of \$0.2billion at the global institution's disposal, shown in the above diagram. According to this reward payment function, a country spending 0.06% of its GDP on public clean energy RD&D and with a GDP making up 1% of the aggregate GDP of all the countries would receive a reward payment of $0.01 \times \$1.8\,billion = \$18 million$.

Zooming out to larger budgets, we find:

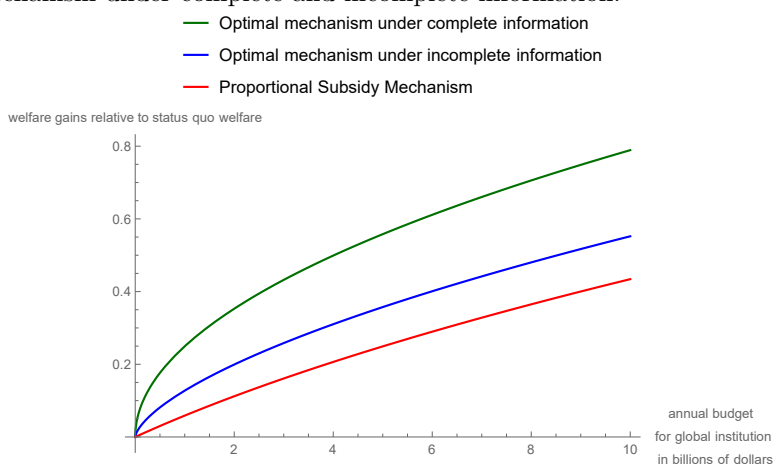


For annual budgets up to $\$10.274\,billion$ there is only partial participation. For annual budgets above this value, there is always full participation and the reward payment only start becoming positive for spending rates at some point strictly (albeit only slightly) larger than the lowest type:



7.3 Welfare gains

As explained in section 3, the global institution takes "global welfare" to be the aggregate rate of research progress. The following plot shows on the y-axis the ratio of the welfare gains achieved due to the mechanism divided by the welfare in the absence of any mechanism. I compute this for both the optimal mechanism under complete and incomplete information:

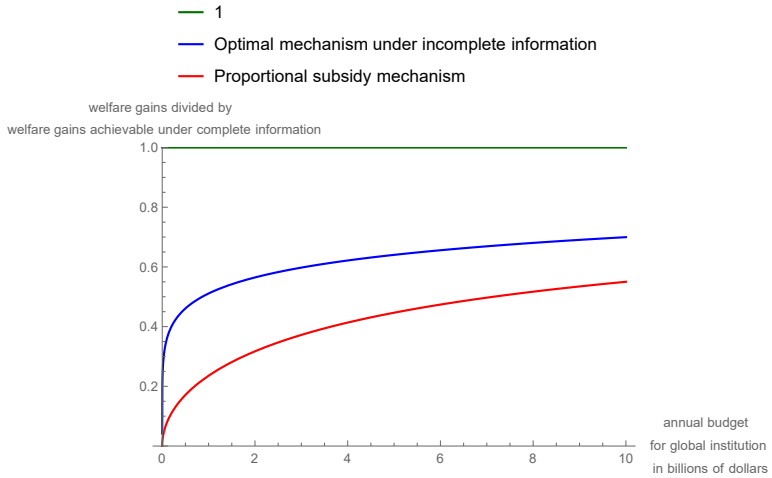


7.3.1 Comparing the optimal mechanism under incomplete information to a Proportional Subsidy Mechanism

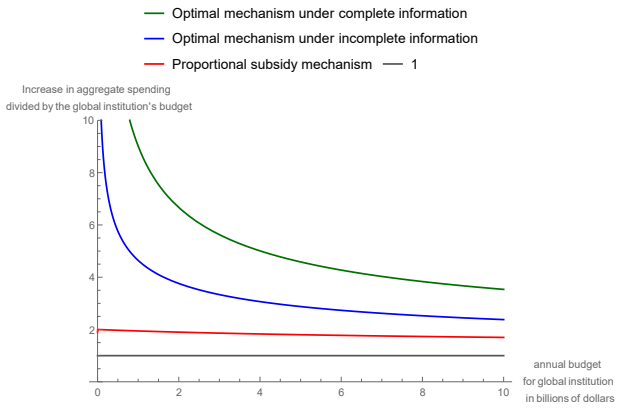
A natural alternative mechanism to consider is a "Proportional Subsidy Mechanism", by which I mean the mechanism that consist of reimbursing countries a fixed percentage of their expenditure on public clean energy RD&D. Under this mechanism, the percentage at which countries are reimbursed is set so as to achieve that the global institution spends exactly its budget.

Arguably, the Proportional Subsidy Mechanism has the important advantage of being very simple and easy to understand. This advantage of simplicity has to be weighed against the lower performance. In the following graph I compare both the Proportional Subsidy Mechanism and the Optimal Anonymous

Mechanism against the optimal mechanism under complete information:

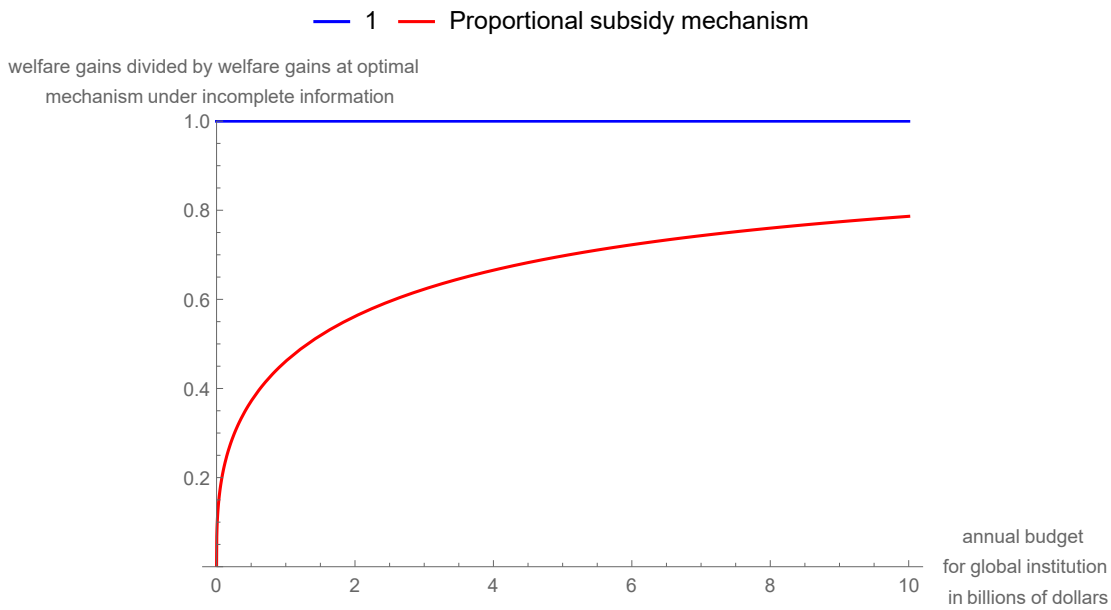


The optimal mechanism under incomplete information substantially outperforms the Proportional Subsidy Mechanism because it manages to reduce informational rents. As a result, it achieves much greater "crowding in" of countries' spending. The following graphs plots an "average crowding-in multiplier", defined as the ratio of the increase in the aggregate spending divided by the global institution's budget:



The size of the crowding-in multiplier shown in this graph highlights the importance of using reward payment mechanisms rather than having the global institution directly fund clean energy research. In fact, under the latter approach we should expect crowding out to occur: governments would have incentives to lower their spending on clean energy RD&D.

Now let us directly compare the Proportional Subsidy Mechanism with the optimal anonymous mechanism. When the global institution's annual budget is \$10 billion dollars, the Proportional subsidy mechanism achieves 83.8% of the welfare gains that the optimal anonymous mechanism achieves. However, for lower budgets the relative performance gap is substantially larger:



At an annual budget of \$2billion, the Proportional Subsidy Mechanism achieves only 63.5% of the welfare gains that the optimal mechanism under incomplete information achieves and at \$1billion only 53.0%.

8 Which mechanism should best be used in practice?

The reward payment functions corresponding to the optimal mechanism under incomplete information, characterized in Lemma 5 and displayed in section 7.2 could be used in practice with a slight modification to ensure budget balance which I will explain below. Such a modification is required to ensure budget balance given that countries are unlikely to react to the incentives exactly in the way predicted by the model.

A natural way to adapt the mechanism to ensure budget balance would be to announce at the beginning of each year the function t rewarding countries for research spending in that year. This function t would be the one that is computed to be optimal within the model. At the end of the year, the actual reward payments would then be set equal to ct, where c is the unique constant that results in budget balance.

Now let us consider alternative mechanisms. There are at least 2 dimensions that one might hope to improve upon by constructing alternative mechanisms: simplicity and performance.

Consider first simplicity. The optimal anonymous mechanism has the drawback of not having an explict formula. In contrast to that, the Proportional Subsidy Mechanism analyzed in section 7.3.1 is maximally simple: Countries

are reimbursed a fixed fraction of their spending on clean energy RD&D, with the value of the fraction being determined by budget balance.

However, in section 7.3.1 we saw that the Proportional Subsidy Mechanism performs substantially worse than the optimal anonymous mechanism: At an annual budget of $2billion, the Proportional Subsidy Mechanism achieves only 63.5% of the welfare gains that the optimal anonymous mechanism achieves and for lower budgets this percentage is even lower. It seems implausible that it would be worth giving up so much welfare for the gain in simplicity.

Now let us consider the second dimension in which we might hope to improve upon the optimal anonymous mechanism, namely performance. To achieve greater welfare gains for a given budget, alternative mechanisms could condition on other variables than just the current spending on clean energy RD&D (and GDP). For example, they could also condition the reward payment to a country on its past spending on clean energy RD&D at times before the introduction of the mechanism. This would allow the mechanism to reward countries for increases in their spending relative to their spending levels in the past. Thereby, informational rents could be reduced and thus greater welfare gains could be achieved.

A potential drawback of this approach could be that it might set a bad precedent. In the future, in the years leading up to the creation of new global institutions rewarding countries for globally beneficial actions countries might (rationally) believe that by contributing less then they will in the future receive larger reward payments.[4]

These perverse incentive effects are potentially large, as I find in Stern (2021). This weighs in favor of a norm requiring global institutions to abstain from conditioning their reward payments on a country's past actions.

Given these drawbacks of conditioning on countries' past actions, it is important to know by how much doing so could increase the immediate welfare gains that the mechanism could achieve. The results in section 7.3.1 provide an upper bound on these welfare gains: If the global institution's annual budget is $5billions then the optimal mechanism under complete information achieves 47.6% more welfare gains than the optimal anonymous mechanism. For larger budgets this potential for improvement declines, reaching 35.7% for an annual budget of $10billion.

In reality, the welfare gains attainable from conditioning on countries' actions prior to the introduction of the mechanism are likely to be much lower. One reason for this is that countries' types might change over time. In early years, a mechanism might be well taylored to the countries' types as they were revealed in the years leading up to the establishment of the mechanism. But later on, this might be less and less the case. Under the interpretation of the model

[4]For example, there are discussions about the establishment of a Global Health Security Challenge Fund that would reward countries on the basis of their scores in the Global Health Security Index (Nuzzo (2021)). Mechanism like these that aim to induce countries to internalise more of the externalities from preventing pandemics will likely become increasingly important in the decades ahead when advances in synthetic biology will increase the destructive potential of pandemics.

presented in this paper, a country's "absorptive capacity for public spending on clean energy RD&D" will evolve over time. As a result, the countries' types will change.

9 Limitations and further research questions

This paper's analysis has restricted attention to mechanisms that reward countries on the basis of their aggregate spending on clean energy RD&D. An alternative approach would be to create a separate global institution for each of the 7 categories of clean energy RD&D that countries report their spending for in the Mission Innovation Initiative. The model presented in the current paper could be applied to each of the 7 categories separately to answer the following question: With a given overall budget and restricting to anonymous mechanisms, which of the following can generate larger welfare gains: The mechanism conditioning reward payments on aggregate research spending on the 7 categories or an optimal combination of separate mechanism, each conditioning reward payments on countries' spending on one of the 7 categories.

A further importan research direction concerns how the mechanism should be updated over time. The current paper focuses on the initial design question: How should a global institution best reward countries for their research spending when it is first established. However, once such an institution is established, countries' observed reactions to the incentives provides valuable information that should be incorporated into the procedure for updating the reward payment functions. For example, this data might allow for a better estimation of the parameter λ and also for more flexible functional forms for the functions translating research spending into progress.

References

[1] Acemoglu, D., Akcigit, U., Hanley, D., & Kerr, W. (2016). Transition to clean technology. Journal of Political Economy, 124(1), 52-104.

[2] Cunliff (2019). Omission Innovation 2.0: Diagnosing the Global Clean Energy Innovation System

[3] Kornek, U., & Edenhofer, O. (2020). The strategic dimension of financing global public goods. European Economic Review, 127, 103423.

[4] Laffont, J. J., & Martimort, D. (2009). The theory of incentives. Princeton university press.

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

[6] Martimort, D., & Sand-Zantman, W. (2013). Solving the global warming problem: beyond markets, simple mechanisms may help!. Canadian Journal of Economics/Revue canadienne d'économique, 46(2), 361-378.

[7] Molina, C., Akçay, E., Dieckmann, U., Levin, S. A., & Rovenskaya, E. A. (2020). Combating climate change with matching-commitment agreements. Scientific reports, 10(1), 1-12.

[8] Myslikova, Z., & Gallagher, K. S. (2020). Mission Innovation is mission critical. Nature Energy, 5(10), 732-734.

[9] Nuzzo, J. (2021). To Stop a Pandemic: A Better Approach to Global Health Security. Foreign Aff., 100, 36.

[10] Sandler, T. (2001). On financing global and international public goods. Washington, DC: World Bank.

A Proofs for section 6

A.1 Proof of Lemma 3

Proof. Incentive compatibility implies:
$\frac{b}{\lambda}(\frac{x(b)}{\lambda})^{\lambda} - \frac{b}{\lambda} + b - x(b) + t(x(b)) \geq \frac{b}{\lambda}(\frac{x(b')}{\lambda})^{\lambda} - \frac{b}{\lambda} + b - x(b') + t(x(b'))$
i.e.:
$\frac{b}{\lambda}(\frac{x(b)}{\lambda})^{\lambda} - x(b) + t(x(b)) \geq \frac{b}{\lambda}(\frac{x(b')}{\lambda})^{\lambda} - x(b') + t(x(b'))$
and symmetrically:
$\frac{b'}{\lambda}(\frac{x(b')}{\lambda})^{\lambda} - x(b') + t(x(b')) \geq \frac{b'}{\lambda}(\frac{x(b)}{\lambda})^{\lambda} - x(b) + t(x(b))$
Adding these inequalities together gives:
$(\frac{b}{\lambda} - \frac{b'}{\lambda})(\frac{x(b)}{\lambda})^{\lambda} \geq (\frac{b}{\lambda} - \frac{b'}{\lambda})(\frac{x(b')}{\lambda})^{\lambda}$
Equivalently:
$(\frac{b}{\lambda} - \frac{b'}{\lambda})((\frac{x(b)}{\lambda})^{\lambda} - (\frac{x(b')}{\lambda})^{\lambda}) \geq 0$
Hence $b > b'$ implies $x(b) \geq x(b')$. □

A.2 Proof of Lemma 4

Proof: First I will show that there exists a number $b_{min} \in [\underline{b}, \bar{b}]$ such that the set of participating types is given exactly by $[b_{min}, \bar{b}]$. To establish this fact, it is convenient to formulate the global institution's problem as a control problem with the state variable $U(b)$ and the control variable $x(b)$:

Problem 3.

$$\max_{x(.)} \int_{b=\underline{b}}^{\bar{b}} f(b) \frac{b^{1-\lambda}}{\lambda} x(b)^{\lambda} db$$

under the following constraints:

$$\int_{b=\underline{b}}^{\bar{b}} (-v(x(b), b) + U(b)) f(b) db \leq M$$

$$U'(b) = \frac{1-\lambda}{\lambda}\frac{x^\lambda - b^\lambda}{b^\lambda} \tag{1}$$

$$U(b) \geq 0$$

$$x(b)\ non-decreasing$$

We will ignore the constraint that $x(b)$ be non-decreasing. It turns out that it will be automatically satisfied. Let $\phi(b)$ denote the Lagrange multiplier associated with the constraint $U(b) \geq 0$.

The Hamiltonian H is given as follows:

$$H = f(b)(\frac{b^{1-\lambda}}{\lambda}x(b)^\lambda - \sigma(-v(x(b),b) + U(b))) + \phi(b)U(b) + \psi(b)\frac{1-\lambda}{\lambda}\frac{x^\lambda - b^\lambda}{b^\lambda}$$

The evolution of the costate variable ψ corresponding to the state variable U is governed by the following condition:

$$\psi'(b) = -\frac{\partial H}{\partial U} = \sigma f(b) - \phi(b) \tag{2}$$

with free endpoint condition:

$$\psi(\bar{b}) = 0$$

The first order condition for the optimality of the control variable x is given by:

$$\frac{\partial H}{\partial x} = f(b)((\frac{b}{x})^{1-\lambda} + \sigma((\frac{b}{x})^{1-\lambda} - 1)) + \psi(b)\frac{1-\lambda}{b}(\frac{b}{x})^{1-\lambda} = 0 \tag{3}$$

Consider any point $\hat{b}$ where the constraint $U(b) \geq 0$ is binding, meaning that $\phi(\hat{b}) > 0$. (There must be such a point, given that the global institution's objective function is unbounded.) Now take the maximal (in the sense of set inclusion) interval containing $\hat{b}$ such that $U(b) = 0$ on the interval. Since the pre-image $U^{-1}(\{0\})$ of the closed set $\{0\}$ must be closed by the continuity of U, this maximal interval must be closed. We therefore denote it by $[b_1, b_2]$. Now the claim of the Lemma follows if we can show that $b_1 = \underline{b}$. Therefore, for a contradiction, suppose that $b_1 > \underline{b}$.

By the definition of the interval $[b_1, b_2]$, there must be points $b \in [\underline{b}, b_1)$ arbitrarily close to b_1 with $U(b) > 0$. By the complementary slackness condition, for these points we must have $\phi(b) = 0$. Hence the continuity of ϕ implies that $\phi(b_1) = 0$.

By equation 1, we must also have $x(b) = b \forall b \in [b_1, b_2]$. Using this in the optimality condition, equation 3, yields:

$$f(b)(1+\sigma) + \psi(b)\frac{1-\lambda}{b} = 0 \forall b \in [b_1, b_2]$$

$$\psi(b) = -\frac{b}{1-\lambda} f(b) \forall b \in [b_1, b_2] \tag{4}$$

From the costate evolution equation 2 we get:

$$\phi(b) = \sigma f(b) - \psi'(b)$$

Substituting in the differentiated equation 4 yields:

$$\phi(b) = (\sigma + \frac{1+\sigma}{1-\lambda}) f(b) + (1+\sigma)\frac{b}{1-\lambda} f'(b) \forall b \in [b_1, b_2]$$

$$\phi(b) = \frac{f(b)}{1-\lambda}(((1-\lambda)\sigma + 1 + \sigma) + (1+\sigma)\frac{bf'(b)}{f(b)}) \forall b \in [b_1, b_2]$$

Now if $\frac{bf'(b)}{f(b)}$ is non-increasing then we must have $\phi(b) \leq \phi(b_1) = 0 \forall b \in [b_1, b_2]$. But this contradicts the fact that $\hat{b} \in [b_1, b_2]$ and $\phi(\hat{b}) > 0$. Thus we have shown that the assumption that $b_1 > \underline{b}$ leads to a contradiction. Hence we must have $b_1 = \underline{b}$.

Let us recap: We have shown the following about the optimal allocation under incomplete information: if the non-negativity constaint $U \geq 0$ is binding at some $\hat{b}$ then the maximal interval containing $\hat{b}$ such that $U(b) = 0$ holds on the interval is of the form $[\underline{b}, b_2]$. Given that global institution's objective is unbounded, its budget constraint must be binding and thus the nonnegativity contraint $U(b) \geq 0$ must be binding for some type. Thus there exists a number $b_{min} \in [\underline{b}, \bar{b}]$ such that the set of participating types is given exactly by $[b_{min}, \bar{b}]$.

Also, we note that for the log-uniform distribution we have $\frac{bf'(b)}{f(b)} = -1$, so the Lemma applies.

In the above argument we assumed that $\frac{bf'(b)}{f(b)}$ is non-increasing. Actually, the argument can be extended to much more general cases, as I will now demonstrate.

For this, first note that we must have $b_2 < \bar{b}$. To see why, suppose that the types $[\bar{b} - db, \bar{b}]$ do not participate for some small db. Now let us construct the following small perturbation: Let us iduce a type $\bar{b} - \hat{d}b \in [\bar{b} - db, \bar{b}]$ to increase their spending on clean energy RD&D by the amount $\hat{d}b$ relative to what they would do in the absence of the mechanism. This creates welfare benefits of order db^2, whilst the types' aggregate disutility from the induced change in action is of order db^3 and the resulting increase in aggregate rents is also of order db^3. Hence this perturbation must increase welfare for sufficiently small values of db.

It follows that there are points $b \in (b_2, \bar{b}]$ arbitrarily close to b_2 with $U(b) > 0$ and hence $\phi(b) = 0$ by the complementary slackness condition. By continuity of ϕ, it follows that $\phi(b_2) = 0$. As established above, we then have:

$$\phi(b) = \frac{f(b)}{1-\lambda}(((1-\lambda)\sigma + 1) + \frac{bf'(b)}{f(b)}) \forall b \in [b_1, b_2]$$

Now suppose again that $b_1 > \underline{b}$. As shown above, we must then have $\phi(b_1) = 0$. Recall that we also have $\hat{b} \in [b_1, b_2]$ with $\phi(\hat{b}) > 0$. So the only way this can happen is for $\frac{bf'(b)}{f(b)}$ to first increase to the right of b_1 and then to decrease to the right of b_2. $\square$

A.3 Proof of Lemma 5

Proof:

Let L denote the global institution's Lagrangian and μ the Lagrange multiplier associated with the budget constraint:

$$L = \int_{b=\underline{b}}^{\bar{b}} f(b)(\frac{b^{1-\lambda}}{\lambda}x^{\lambda} - \mu(-v(x(b), b) + U(b)))db \tag{5}$$

Using integration by parts, we get:

$$\int_{b=\underline{b}}^{\bar{b}} f(b)U(b)db = [-(1 - F(b))U(b)]_{b=\underline{b}}^{\bar{b}} + \int_{b=\underline{b}}^{\bar{b}} (1 - F(b))U'(b)db$$

$$\int_{b=\underline{b}}^{\bar{b}} f(b)U(b)db = U(\underline{b}) + \int_{b=\underline{b}}^{\bar{b}} (1 - F(b))U'(b)db$$

Now using that $U'(b) = \frac{1-\lambda}{\lambda b^{\lambda}}(x^{\lambda} - b^{\lambda})$, we get:

$$\int_{b=\underline{b}}^{\bar{b}} f(b)U(b)db = U(\underline{b}) + \frac{1-\lambda}{\lambda}\int_{b=\underline{b}}^{\bar{b}} (1 - F(b))\frac{x^{\lambda} - b^{\lambda}}{b^{\lambda}}db$$

Substituting this into equation 5 yields:

$$L = \int_{b=\underline{b}}^{\bar{b}} f(b)(\frac{b^{1-\lambda}}{\lambda}x^{\lambda} - \mu(-v(x(b), b) + max(0, \frac{1-\lambda}{\lambda}\frac{1-F(b)}{f(b)}\frac{x^{\lambda} - b^{\lambda}}{b^{\lambda}})))db \tag{6}$$

Substituting in the fact that $-v(x(b), b) = x - b + \frac{b}{\lambda}(1 - (\frac{x}{b})^{\lambda})$ yields:

$$L = \int_{b=\underline{b}}^{\bar{b}} f(b)(\frac{b^{1-\lambda}}{\lambda}x^{\lambda} - \mu(x - b + \frac{b}{\lambda}(1 - (\frac{x}{b})^{\lambda}) + max(0, \frac{1-\lambda}{\lambda}\frac{1-F(b)}{f(b)}\frac{x^{\lambda} - b^{\lambda}}{b^{\lambda}})))db \tag{7}$$

The first order condition for $x(b)$ if b participates is thus:

$$(\frac{b}{x})^{1-\lambda} - \mu(1 - (\frac{x}{b})^{\lambda-1} + \frac{1-\lambda}{b}\frac{1-F(b)}{f(b)}(\frac{x}{b})^{\lambda-1}) = 0$$

$$(\frac{x}{b})^{1-\lambda} = \frac{1 + \mu - \mu\frac{1-\lambda}{b}\frac{1-F(b)}{f(b)}}{\mu}$$

$$x = b(\frac{1+\mu-\mu\frac{1-\lambda}{b}\frac{1-F(b)}{f(b)}}{\mu})^{\frac{1}{1-\lambda}}$$

□

A.4 Proof of Lemma 6

Proof. For $b \geq b_{min}$ we have:

$$U(b) = \int_{z=b_{min}}^{b} U'(z)dz = \int_{z=b_{min}}^{b} \frac{\partial v}{\partial b}(x(z), z)dz$$

We also have the identity:

$$U(b) = v(\tilde{x}(b), b) + t(\tilde{x}(b))$$

Combining these yields:

$$t(\tilde{x}(b)) = \int_{z=b_{min}}^{b} \frac{\partial v}{\partial b}(x(z), z)dz - v(\tilde{x}(b), b)$$

Now given that the optimal allocation $\tilde{x}(b)$ is strictly increasing on $[b_{min}, \bar{b}]$, its restriction to this interval, $\tilde{x}|_{[b_{min},\bar{b}]}$, is invertible. We thus obtain for $x \in Image(\tilde{x}|_{[b_{min},\bar{b}]})$ by substituting in $\tilde{x}^{-1}(x)$ in the previous equation:

$$t(x) = \int_{z=b_{min}}^{\tilde{x}^{-1}(x)} \frac{\partial v}{\partial b}(\tilde{x}(z), z)dz - v(x, \tilde{x}^{-1}(x))$$

□

B Mathematica notebooks

The Mathematica notebook that contains the data and computes the continuous type distribution from it is available here.

The Mathematica notebook containing all the computations and plots is available here.