Tradable Climate Liabilities: A Thought Experiment

ÉTIENNE **BILLETTE DE VILLEMEUR**

JUSTIN **LEROUX**

2018S-43
WORKING PAPER

CS

CIRANO
Allier savoir et décision

2018s-43

Tradable Climate Liabilities: A Thought Experiment

Étienne Billette de Villemeur, Justin Leroux

Série Scientifique
Scientific Series

Montréal
Décembre/Decembre 2018

CIRANO
Allier savoir et décision

Centre interuniversitaire de recherche en analyse des organisations

ISSN 2292-0838 (en ligne)

Tradable Climate Liabilities: A Thought Experiment [*]

Étienne Billette de Villemeur [†], Justin Leroux [‡]

Abstract

We envision the creation of a climate liability market to address climate change. Each period, countries are issued liability commensurate to their emissions of the period. Liability bearers are required to pay over time, as climate harm materializes. Revenues are used to compensate participating countries in proportion of climate harm. Because liabilities are traded like financial debt among participants, the mechanism achieves a unique carbon price through decentralization of the choice of a discount rate as well as beliefs about the severity of the climate problem. We discuss properties of such a mechanism along the dimensions of efficiency, fairness, exposure to risk, commitment, participation, as well as implementation challenges.

Keywords: Climate Liability, Market Instruments, Pigovian Tax, Risk Sharing

JEL Codes: Q54, H23

---

[*] The authors wish to thank Geir Asheim, Hassan Benchekroun, Yann Bramoullé, Alessio D'Amato, Julien Daubanes, Georges Dionne, Paul Ekins, Gérard Gaudet, Andre´ Grimaud, Christian Gollier, Jérémy Laurent-Luchetti, Peter Hartley, Marcus Hermann, Louis Hotte, Shi-Ling Hsu, Yann Kervinio, Sonja Köke, Stéphane Lambrecht, Derek Lemoine, Virginie Maris, Leslie Martin, Kian Mintz-Woo, Christian de Perthuis, Stephen Salant, Marc San-tugini, Charles Séguin, Steven Shavell, Jens Steffek, Nicolas Treich, Edilio Valentini, Bruno Versaevel, Mirko Wiederholt, Ralph Winkler as well as participants of the Occasional California Workshop on Environmental and Resource Economics, the Montreal Resource and Environmental Economics Workshop, the Paris Environmental and Energy Economics Seminar (PEEES), and the GERAD Conference on Dynamic Games. Leroux is grateful for funding from FRQSC.

[†] LEM UMR 9221, Université de Lille, France.

[‡] HEC Montréal, CIRANO & CRÉ. 3000 chemin de la Côte-Sainte-Catherine, Montreal, QC H3T 2A7, Canada. E-mail: justin.leroux@hec.ca. Corresponding author.

1 Introduction

Climate change is already impacting populations. Insurance markets are responding accordingly: In 2016, economic losses attributed to weather-related events worldwide totaled USD 123 billion, of which USD 37 billion were insured (Swiss Re Institute, 2017). Given the amounts involved, responsibility for climate harm is no longer a question of mere abstract justice, but a matter of very real consequences. Moreover, there seems to be a mounting willingness to adress "loss and damage associated with the adverse effects of climate change" (UNFCCC, 2015, Art. 8) as outlined in the Paris Agreement : At least one company is being sued over its responsibility in current climate harm due to historical emissions (Frank, 2017) and evidence has been compiled tracing more than two thirds of cumulative greenhouse gas (GHG) emissions to existing businesses (CDP, 2017).

The aim of this article is to envision how this recent shift in attitudes, towards assessing the realization of climate harm and holding emitters accountable, could translate into an economic instrument. The mechanism we envision works as follows: when a country emits GHGs at any given date, it contracts a financial debt of sorts, that will have to be repaid over time as climate harm materializes. The funds collected go towards compensating the climate harm of participating countries. This *climate liability* is a tradable financial title, that can be exchanged among participating countries. Whether a country chooses to purchase or to sell liabilities depends on its expected discounted value of global marginal harm. Introducing a market for climate liability allows for the aggregation of these values across countries into a single carbon price.

The mechanism operates in three steps, to be repeated at regular intervals (say, every 5 years): 1) Evaluating the climate harm suffered by each country for the current period in excess of a baseline (say, climate harm circa 1990). Denote by h^i the excess harm suffered by country i, adding up to H globally; and 2) Charging each country a price λ per unit of its contribution to the increase in the stock of GHGs since the baseline year (i.e., 1990 in our example), S^i, so that each country is required to pay $S^i\lambda$. 3) Redistributing the proceeds to countries in proportion of the excess climate harm they suffer.

The mechanism adopts an *ex post* standpoint rather than the familiar *ex ante* approach of all economic instruments introduced thus far, from carbon taxes to cap-and-trade schemes.[1] This article is devoted to exploring the pros and the

[1] It should be clear to the reader that although we use the term "liability", we do not

cons of this reversed perspective. The analysis will focus on the specific liability scheme where liability payments are set equal to marginal harm: $\lambda \equiv H'$.

Although motivated chiefly by compensation (Art. 8 of the Paris Agreement), it turns out that setting liability payments equal to marginal harm incentivizes emissions reductions in an efficient way (Art. 6 of the Paris Agreement): when countries are homogeneous in how they treat the future (discounting) and forecast climate harm (expectations), the mechanism leads to emissions patterns that are first-best efficient (Proposition 1). Furthermore, if countries are homogenous, there is no need for countries to trade liabilities.

If countries disagree in their calculation of the expected discounted harm, then the existence of a market for liabilities becomes useful. Importantly, it allows for the possibility of arriving at an efficient outcome without the need of parties reaching an agreement on the discount rate nor on future climate scenarios. Agreement is reached through decentralization. We show that when default risk is taken into account, the unique (shadow) price on GHG emissions is higher than the social cost of carbon (Proposition 2), thus leading to a further reduction in emissions.

We also quantify the temptation for any country to default on its liability payments and show that it is no greater than the temptation to delay the implementation of a Pigovian carbon tax (Proposition 3). These results indicate that the political economy problems surrounding an ex post approach are just as severe—no more, no less—than under the familiar ex ante viewpoint.

Yet, to be truly effective, the mechanism must garner (voluntary) participation. Because the mechanism operates transfers that are aimed at compensating harm rather than garanteeing participation, some net contributors may not see participation as an improvement of their own condition over the business-as-usual scenario. This is the price to pay to correct for the inequalities generated by the spatial idiosyncracies of climate harm. Nevertheless, if participants are similar enough, the liability scheme results in a Pareto improvement due to the fact that the joint gains from reduced climate harm more than makes up for the outlays of the net contributors (Proposition 4). Moreover, the more participants, the higher the incentives to participate; this hints at the possibility of a positive unraveling effect spurred by an initial "coalition of the willing".

The formal model and the above results are presented in Section 3, after an

intend that the climate change problem is to be solved by resorting to legal action—i.e., court trials—to punish tortfeasors (which are essentially every single country, because all bear some responsibility for climate harm for ever having emitted greenhouse gases).

overview of the related literature (Section 2). The remainder of the article is a discussion of the features of the mechanism as informed by the model (Section 4) as well as a broader discussion, extending beyond the model, of the implementation challenges of a climate liability market and of the commitment and participation (Section 5).

In Section 6, we will reach the conclusion that choosing between an ex ante or an ex post policy amounts to choosing between, on the one hand, an absence of a (long-term) commitment—with the ex ante approach—and, on the other hand, a genuine responsibility towards the consequences of climate change, greater international participation, and a more equitable distribution of the burden across generations—with the ex post approach. Economic efficiency, however, is unaffected by the choice of viewpoint. In fact, this thought experiment leads us to the realization that a climate policy can follow the Pigovian logic and yield first-best incentives while not exacting the full payment of the social cost of carbon at the time of emission.

2 Relation to the literature

The idea of using liability as a means to control externalities traces back to Calabresi (1970) and was recently compared to corrective taxation in Shavell (2011).[2] On the one hand, regulation (i.e., taxation) is costly even in the absence of harm, whereas liability only kicks in when harm actually occurs. On the other hand, a liability approach is typically more informationally demanding because, following legal procedure, it requires establishing tort (Kolstad et al, 1990; Shavell, 2011) or at least being able to observe the level of precautionary effort exerted by the examined party (Hiriart et al., 2004, 2010).[3] Hence, liability is likely to be more appropriate in situations where harm is highly uncertain but where its source can be easily established. This is precisely the case of climate change, where the magnitude of harm is typically unknown ex ante but the responsibility of countries towards CO_2 concentration can be readily established thanks to available data on cumulative CO_2 emissions per country

[2] The issue of liability for climate change has already been raised for some years by legal scholars (University of Pennsylvania Law Review, 2007; Faure and Peeters, 2011; Lord et al., 2012), who focus on the legal framework's ability to handle climate-related losses.

[3] The liability approach is usually discussed in the context of tort law, involving private parties and legal costs attached to lawsuits, to establishing due care and negligence. By contrast, the liability approach we consider here is public, in the sense that it involves countries, and would consist in an automatic procedure where the negligence rule plays no role.

(e.g., from the World Resource Institute or the World Bank databases). The general argument echoes that of Shleifer (2012), according to which the need for regulation arises where litigation is ineffective. Underlying this line of reasoning is the notion that turning to litigation—i.e., liability—is a most natural reflex that should be left unhindered whenever it is an efficient option.

The use of liabilities to address the climate problem is further supported by insights from the cost-sharing literature. An important lesson to be learned from that literature is that the best properties of a payment scheme—whether in terms of efficiency, incentives, and even fairness—arise from mirroring the physical features of the externality to be managed (Moulin, 2002). Along those lines, and as we shall evidence below, the very fact of asking payments as long as damage occurs (instead of ex ante) is likely to result in a strict improvement along several dimensions. Climate harm being a problem caused by the *stock* of CO_2 in the atmosphere, it is thus appropriate to condition payment on emission stocks rather than on emission flows as do carbon taxes and cap-and-trade programs. Ambec and Ehlers (2016) have recently shown in a wide variety of externality problems that the polluter-pays principle (i.e., requiring polluters to compensate victims) leads to the unique welfare-enhancing redistribution that also holds polluters responsible for their actions. Carbon liabilities can be seen as expanding on the polluter-pays principle by focusing on specific features of the climate externality problem, particularly by unfolding the time and uncertainty dimensions. By making explicit—and, most importantly, financial—the somewhat intangible carbon debt that mankind accumulates along with atmospheric CO_2, climate liabilities do just that.

In a climate change context, liability has very recently been proposed as a means of making global cooperation more effective and less costly in the long run (Gampfer, Gsottbauer and Delas, 2014). Its merit lies in the fact that countries are more likely to adhere to an agreement on emissions reductions if they believe they will be compensated fairly for future climate damage incurred (Gampfer, 2014). As in Billette de Villemeur and Leroux (2011), the argument in these works is one of fairness. Although we include a fairness objective (i.e., compensating the idiosyncrasies of climate harm), an important feature of the liability mechanism is that it comes at no cost to efficiency. Moreover, the mechanism makes participation appealing (see Proposition 4 and Section 5.3 below). This lessens the distortions that arise due to limited participation. In other words, climate liabilities mitigate the externality on a larger scale (more participants) and garner participation without compromising efficiency.

Our contribution is related to the literature on green accounting, which incorporates environmental externalities into national accounting (see Weitzman 1976; Hartwick, 1990; or, more recently, Cairns, 2004, Cairns and Lasserre, 2006). The goal there is to give an intertemporal account of a country's available resources so as to measure how much is left to future generations. Many indicators, environmental and otherwise, including more than 150 Sustainable Development Goal Indicators proposed formally by the United Nations Economic and Social Council (UN Statistical Commission, 2016), have flourished over the past decades. While informative, they are only tenuously tied to precise action, if at all. This calls into question the incentives they are able to generate. By contrast, the liability mechanism consists in providing a single indicator of indebtedness for climate change that directly impacts policy, because climate liabilities imply payments upon the realization of harm. Moreover, the financial stakes involved generate powerful incentives to provide accurate predictions of future harm. As such, our work is in line with the burgeoning literature on stakeholder value. The latter argues in favor of moving away from the conventional shareholder value maximization to *stake*holder value maximization (Magill et al., 2015). In our proposal, the financialization of carbon emissions as debt can be seen as giving voice to global (and future) stakeholders.

Many authors have proposed alternatives to a carbon tax and to cap-and-trade programs with the aim of facilitating global cooperation. Closely related to ours are two proposals by Gersbach and Winkler. Gersbach and Winkler (2011) propose an international emissions permit market with refunding, which they find to lead to welfare improvements and, if the share of freely allocated goes to zero, to a level of emissions that is close to optimal. The liability mechanism we envision achieves exactly the optimal emissions path when countries are identical (Gersbach and Winkler only consider the case of identical countries), with welfare improvements being guaranteed when countries are homogeneous enough (Proposition 4).

At the opposite end of the price-versus-quantities spectrum, Gersbach and Winkler (2012) propose a global refunding scheme that is similar to our liability mechanism in that both experiment with the timing of the payments. In our case, payments are backloaded, incentivizing participation at the expense of commitment. By contrast, their scheme does the opposite: because countries must contribute large sums to a fund at the start (which are then reimbursed as a function of emissions abatement) commitment is not an issue, but participation certainly is. In sum, there is a symmetry of sorts between the liability mechanism we envision and the global refunding scheme of Gersbach and Winkler.

Of course, the effectiveness of the liability scheme rests upon the hypothesis that countries emitting today will have the ability to pay their dues in the distant future. Because emissions reductions can be seen as an investment towards reduced future climate harm, the liability mechanism may seem vulnerable to the maturity mismatch problem. The latter refers to situations whereby the benefits of an investment arrive too late for a firm to be able to repay its debt, or for a bank to honor withdrawals by depositors. It is a serious concern, because the maturity mismatch problem is at the core of the main models of bank runs (Diamond and Dybvig, 1983) and, more generally, of systemic risk to financial and insurance systems (Bobtcheff et al., 2016). Yet, substantial differences between a liability scheme and the banking system lead us to believe that, while deserving attention, the issue is unlikely to put the liability mechanism at risk. First, although countries can choose their exposure to climate risk, they cannot choose the maturity of climate risk because it is exogenous. As a result, the moral hazard issue is less severe than that usually attached to the banking sector (Farhi and Tirole, 2012). Second, the very participation of countries to the liability scheme reduces rather than magnifies the overall magnitude of climate risk. Moreover, exposure to the maturity mismatch problem is less severe under a liability scheme than under, say, the equivalent tax scheme: both generate the same climate benefits while cumulative liability payments are less at any date (Billette de Villemeur and Leroux, 2012).[4] Third, maturity risk is actually foreseeable because liability payments are proportional to marginal harm, which is evaluated on the basis of trends rather than on contemporary realizations of harm. Thus, just like the rest of the (re-)insurance sector, the liability scheme can be considered to be exempt from the systemic risk issue (Kessler, 2014).

This article is related to several works concerned with addressing the intergenerational aspect of environmental policy by transferring income to the earlier generations (who are making mitigation efforts) from the later generations (who enjoy the benefit of a less degraded climate). Formulated in OLG models, the income-transferring devices can take the form of debt financing (Bovenberg and Heijdra, 1998), of an intergenerational trust fund (Gerlagh and Keyzer, 2001),

[4]We adopt the view that failing to honor tax payments is just as serious a breach of responsibility than failing to honor liability payments. Both instruments are a financial manifestation of the environmental debt contracted by emitters.

or of a pension “deal” between young and old (as in PAYGO pension systems; von Below et al., 2016). All find that efficiency and Pareto improvements over the business-as-usual scenario can be achieved. The fact that a liability scheme spreads payments over time is also a step towards relieving the current generation of some of the costs of its effort. Moreover, climate risk is now borne by emitters—rather than by countries suffering climate harm—so that, as compared to ex ante instruments, a liability scheme induces important changes around how risk is shared, both over generations and spatially. In our setting, capital accumulation is exogenous (but unconstrained) to rather account for the heterogeneity of countries. Heterogeneity allows us to give insights about the possibility of decentralizing important parameters (beliefs and discounting) and about the sharing of risk (both spatial and intertemporal).

The ability to decentralize views on discounting is particularly useful because the choice of a discount rate is at the heart of debates in the literature on the social cost of carbon, including the well-known Stern-Nordhaus debate (Nordhaus, 2007) and the Weitzman-Gollier puzzle (Buchholz and Schumacher, 2008; Gollier and Weitzman, 2010). The stakes surrounding these debates are extremely high, which does not make it easy for stakeholders to reach an agreement. By decentralizing the discount rate, a liability scheme attenuates the tensions surrounding negotiations (Section 5.3.2). Moreover, being able to express a country-specific discount rate is all the more useful as inter- and intra-regional inequalities have an impact on the appropriate value of the discount rate (Dennig et al., 2015).

In some sense, our thought experiment parallels that of Weitzman (2017), which suggests that countries should vote on the social cost of carbon along the principle of one-person one-vote, a median-voter result then yielding the appropriate (uniform) global tax rate. Instead, by relying on a market mechanism to aggregate differing opinions, we avoid the bias attached to focusing on median values, which generally do not coincide with optimal values.

3 A market for climate liabilities

Consider a set $N = \{1, ..., n\}$ of countries and denote by e_t^i the flow of emissions of Country i during period t. Emissions are an input in the production function and give rise to a flow of benefits: $b_t^i\left(e_t^i\right)$.

We assume that each country, i, is endowed with time preferences that can be represented by the standard exponential discounting model with its own constant discount factor, $0 < \beta_i < 1$.[5] Thus, each country taken in isolation exhibits a time-consistent emissions plan that is easy to characterize. This is for expositional purposes only, because our results would go through if countries adopted a non-constant discount factor, as the logic would be unaffected (see footnote 10).[6]

Define a reference date, $t = 0$, as a starting point to the mechanism. Emissions have a persistence rate of γ,[7] so that the contribution of Country i to the atmospheric stock of GHGs since the reference date is

$$S_t^i = \sum_{s=0}^{t} \gamma^{t-s} e_s^i - \mathcal{S}_0^i, \tag{1}$$

where $\mathcal{S}_0^i$ is the stock attributable to Country i at the reference date.[8] Each period, the stock of GHGs generates a (stochastic) flow of excess harm $h_t^i(S_t)$ in Country i, relative to the baseline year.[9] We denote by $H_t(S_t) = \sum_i h_t^i(S_t)$ the flow of global excess harm, where $S_t = \sum_{i=1}^n S_t^i$. From now on, we will use the term "harm" to mean "excess harm", for convenience.

Denote by $M \subseteq N$ the coalition of participating countries. To be a member of the coalition M, a country must 1) agree to be a part of it (the mechanism is voluntary) and 2) be in good standing with respect to its climate liability payments (no delinquent climate debt). The financial contributions requested of countries are based upon their contributions to the stock of GHGs, the cause of climate harm. In each period, contributions to the carbon stock are priced at λ_t—possibly set equal to the current global marginal harm, H_t', as will be the

[5] Since Koopmans (1960), it is known that this functional form is the only one that grants independence and stationarity to (individual countries') solutions of dynamic choice problems. By allowing countries to exhibit heterogeneous discount rates, our (multi-agent) problem is equivalent to one with a (single) representative agent endowed with a non-constant discount rate (Weitzman 1998, 2001; Gollier and Zeckhauser, 2005; Heal and Millner, 2014).

[6] Uncertainty about the proper discount rate is another reason why discounting ought to be done at a non-constant (decreasing) rate (Gollier and Weitzman, 2010).

[7] For ease of exposition, we shall assume the decay rate to be constant over time and independent of GHG concentration. In all likelihood, it is not the case. Section 4.1.1 discusses why departures from this assumptions are actually not problematic from an ex post perspective.

[8] It is actually not required to trace back emissions to infinity. In fact, accounting only for, say, post-1990 emissions would result in the very same emissions pattern. The truncation simply amounts to lump-sum transfers to larger historical emitters while preserving incentives at the margin, as we discuss in Section 5.3.6.

[9] Formally, excess harm is obtained by subtracting from the observed damage at date t, $d_t^i(S_t + \mathcal{S}_0)$, a damage value, $d_0^i(\mathcal{S}_0)$, associated to the reference date: $h_t^i(S_t) = d_t^i(S_t + \mathcal{S}_0) - d_0^i(\mathcal{S}_0)$, where $\mathcal{S}_0 = \sum_{i=1}^n \mathcal{S}_0^i$. We assume that the h_t^i's are positive and increasing in their argument.

focus of our analysis—so that Country $i \in M$ is charged $\lambda_t S_t^i$ as harm occurs. Proceeds are then distributed in proportion to harm among the participating countries: Country $i \in M$ receives $\left(\lambda_t \sum_{j\in M} S_t^j\right) \left[h_t^i\left(S_t\right) / \sum_{j\in M} h_t^j\left(S_t\right)\right]$ as compensation.

Note that per capita considerations are absent of our analysis, both on the liability side and on the compensation side. Regarding liability, although demographic changes may influence the effective impact of liability payments on individuals over time, our focus is on efficiency. This requires the mechanism to apply directly to the emissions of countries, regardless their source (and, also, regardless of demography). Moreover, regarding compensation, if one makes the first-order approximation that harm is proportional to population, compensating countries in proportion to harm amounts to accounting both for heterogeneity in population and in per capita harm.

Consider now the incentives generated by this mechanism. Assuming that the coalition M is stable over time, Country i's objective function is to maximize, at each date t, the expected discounted sum of its payoffs:

$$\max_{\{e_u^i\}_{u=t}^{+\infty}} E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \left[b_u^i\left(e_u^i\right) - h_u^i\left(S_u\right) - \lambda_u S_u^i + \lambda_u \left(\sum_{j\in M} S_u^j\right) \frac{h_u^i\left(S_u\right)}{\sum_{j\in M} h_u^j\left(S_u\right)} \right] \right\}, \tag{2}$$

where E_t^i denotes the expectation by Country i at date t.

Country i's optimal emissions pattern is then given by:

$$\frac{db_t^i}{de_t^i} = \frac{d}{de_t^i} E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \left(h_u^i\left(S_u\right) + \lambda_u \left[S_u^i - \left(\sum_{j\in M} S_u^j\right) \frac{h_u^i\left(S_u\right)}{\sum_{j\in M} h_u^j\left(S_u\right)} \right] \right) \right\} \tag{3}$$

$$= \frac{d}{de_t^i} E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \left(h_u^i\left(S_u\right) + \lambda_u \left[S_u^i \left(1 - \frac{h_u^i\left(S_u\right)}{\sum_{j\in M} h_u^j\left(S_u\right)}\right) - \left(\sum_{j\in M\setminus\{i\}} S_u^j\right) \frac{h_u^i\left(S_u\right)}{\sum_{j\in M} h_u^j\left(S_u\right)} \right] \right) \right\}. \tag{4}$$

In practice, the ratio $h_t^i\left(S_t\right) / \sum_{j\in M} h_t^j\left(S_t\right)$ cannot be affected by any one country's emissions over a single period. Similarly, the emissions of one country in a given period cannot affect the liability rate λ_t, even if it is anchored to marginal harm, $H_t'(S_t)$. It follows that (4) can be rewritten as:

$$\frac{db_t^i}{de_t^i} = E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \left[\frac{\partial h_u^i}{\partial S_u} + \lambda_u \left(1 - \frac{h_u^i(S_u)}{\sum_{j \in M} h_u^j(S_u)} \right) \right] \frac{\partial S_u^i}{\partial e_t^i} \right\}, \tag{5}$$

where $\partial S_u^i / \partial e_t^i = \gamma^{u-t}$.

We shall work under the benchmark assumption that each country's harm is small relative to the global harm and that the size of the coalition M is large enough for redistribution not to impact incentives. (Appendix B shows how incentives are affected when the coalition M is small.) These standard assumptions lead to:[10]

$$\frac{db_t^i}{de_t^i} = E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \gamma^{u-t} \lambda_u \right\}. \tag{6}$$

When setting $\lambda_t \equiv H_t'$, the latter expression becomes:

$$\frac{db_t^i}{de_t^i} = E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \gamma^{u-t} \frac{dH_u}{dS_u} \right\} = \tau_t^i, \tag{7}$$

where τ_t^i is the standard Pigovian tax on emission flows associated with discount factor β_i and the expectations operator E_t^i.[11,12]

If discount factors and expectations operators are identical across countries, the socially optimum emissions path is easy to define. The τ_t^i's are then all equal to the optimal Pigovian tax and the liability mechanism, with $\lambda_t \equiv H_t'$, thus implements the first-best allocation.

When discount factors and/or expectations differ, the social optimum is ill-defined. For instance, heterogeneity in the discount factors necessarily results in time inconsistency of the representative agent (Zuber, 2011; Jackson & Yariv, 2015).[13] Moving away from the representative-agent approach, we argue that

[10] If the discount factor is not time invariant, so that it can be denoted $\beta_{i,t}$, then β_i^{u-t} in Expressions 2-7 should be substituted by $\prod_{s=t}^{u} \beta_{i,s}$. As mentionned in footnote 7, a similar operation can be made with the decay rate γ.

[11] In a different setting, Benchekroun and Long (1998) establishes a similar equivalence between an optimal tax on pollution stock with an optimal tax on flows.

[12] In the symmetric case, as the one considered in Gersbach and Winkler (2011), a liability scheme can lead to optimal incentives as soon as more than one country participates: setting $\lambda_t = |M| H_t' / (|M| - 1)$, leads to Expression (7) because then $\lambda_t \left(1 - \frac{h_t^i(S_t)}{\sum_{j \in M} h_t^j(S_t)} \right) = H_t' \gg \frac{\partial h_t^i}{\partial S_t}$.

[13] Heal & Millner (2018) recently evidenced that the heterogeneity of discount factors does

the mechanism generates proper incentives because it yields allocative efficiency at each period (marginal benefits are equalized at all times); moreover, it does so while granting equal weight to the opinion of each country.

Proposition 1. *Climate liabilities generate proper incentives towards the efficient emissions pattern.*

Clearly, a scheme equivalent to heterogenous τ_t^i's would generate inefficiencies because marginal benefits would not be equal across countries. To harmonize these implicit tax rates, we introduce a *climate liability market*, which operates as follows. Upon emitting GHGs in the atmosphere, Country i is issued 'climate liabilities' commensurate to its emissions, $l_t^i = e_t^i$. Liabilities do not expire, but instead decay at the same rate, γ, as GHGs. They are traded on a world market at price p_t among countries in the participating coalition. Denote by x_t^i the net quantity of liabilities sold by Country i in period t. The outflow of payments[14] associated with the sale of current liabilities is $p_t x_t^i$ while Country i's cumulative liability holdings in period t writes as:

$$L_t^i = \sum_{u=0}^{t} \gamma^{t-u} \left(l_u^i - x_u^i \right). \tag{8}$$

Accordingly, liability bearers are required to pay damages over time as climate harm occurs, so that Country i is now charged $\lambda_t L_t^i$ in period t. Holding climate liability becomes akin to holding debt.

In practice, as debt grows, national economies often incur extra costs due to increased borrowing constraints (Wachtel and Young, 1987; Engen and Hubbard, 2004; Laubach, 2009). We account for this by introducing a convex cost function, C_t^i, associated with holding climate debt.[15]

The objective of Country i becomes:

$$\max_{\{e_u^i, x_u^i\}_{u=t}^{+\infty}} E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \left[b_u^i \left(e_u^i \right) - h_t^i \left(S_t \right) - \lambda_u L_u^i + \lambda_u \sum_{j \in M} S_u^j \frac{h_u^i \left(S_u \right)}{H_u \left(S_u \right)} - p_u x_u^i - C_u^i \left(L_u^i \right) \right] \right\}. \tag{9}$$

not rule out time inconsistency outright, but imposes the equally troublesome requirement that "the welfare weights attached to individual's utilities evolve appropriately with calendar time".

[14] Because the buyer endorses the responsibility for future payments associated to the liability, it will only agree to hold it in exchange of payment *from* the seller *to* the buyer.

[15] The cost of holding debt is country specific so as to reflect the heterogeneity of countries, including in their amount of other, non climate-related financial debt.

This program yields the following first-order condition:

$$e_t^i: \quad \frac{db_t^i}{de_t^i} = \quad E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \gamma^{u-t} \left(\lambda_u + \frac{dC_u^i}{dL_u^i} \right) \right\}, \tag{10}$$

$$x_t^i: \quad p_t = \quad E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \gamma^{u-t} \left(\lambda_u + \frac{dC_u^i}{dL_u^i} \right) \right\}. \tag{11}$$

When both (10) and (11) hold true, the marginal benefit of emissions is equal to the market liability price. In other words:

Proposition 2. *The introduction of a market for liabilities implicitly yields a single global (shadow) price of emissions.*

In essence, climate liabilities act as tradable Arrow-Debreu-type securities that make markets more complete, thus yielding allocative efficiency through decentralization. Countries that expect global harm to be larger—relative to other countries—and that place a higher value on the welfare of future generations wish to get rid of their liability and end up being net sellers. At one extreme, some countries may even sell as much liability as allowed by their emissions ; those countries actually emit GHGs *as if* emissions were priced at a higher rate than the market price.[16] Conversely, countries that expect global harm to be smaller and place a lower value on the welfare of future generations are willing to hold additional liabilities in exchange of monetary compensation. Hence, even countries that are typically reluctant to take part in a climate agreement see a financial advantage in participating in the climate liability mechanism.[17] At the extreme, countries with little regard for climate harm and likely to succumb to short-termism may wish to acquire as much liability as possible in order to reap the immediate lump-sum payments. In principle, such cases should be the exception, because accumulating liabilities is very costly: financial costs rise as liability grows, due to a country's financial rating deteriorating. Moreover, defaulting on payments should be considered with care by countries, owing to the fact that climate liabilities are a financial title like any other: defaulting would

[16]This situation amounts to a corner solution in x_t^i, where

$$\frac{db_t^i}{de_t^i} = E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \gamma^{u-t} \left(\lambda_u + \frac{dC_u^i}{dL_u^i} \right) \right\} > p_t.$$

[17]As mentioned, the mechanism supposes that a country must fully take part in the mechanism—i.e., to be attributed liabilities for its own emissions and to be in good standing with regards to paying the associated damages—in order to be allowed purchase liabilities on the market.

not only exclude the country from receiving compensation transfers—as per the mechanism itself—but may also likely result in exclusion from other financial markets. We discuss additional safety mechanisms to prevent such cases from occuring in Section 5.2.3.

Notice that when $\lambda_t \equiv H_t'$, the equilibrium liability price is greater than the expected discounted marginal harm (Expression 11). This is because of the cost of holding liability, which reflects the default probabilities of countries. At the country level, the higher the default risk, the larger that country's marginal cost of holding liability, and the greater the incentive to get rid of it. At the aggregate level, this puts an upward pressure on the liability price—a seller must offer to pay more to find a willing buyer—which, in turn, further reduces global emissions.

By its ex post nature, the effectiveness of the climate liability scheme relies on the assurance that countries will honor their liability payments over time. If countries were to default on their payments, populations suffering harm would not be compensated. Not only that, but the very *anticipation* of one's own default would severely weaken one's incentives to cut back on emissions.[18] That being said, defaulting is just a form of free riding. From the viewpoint of the individual country, *any* climate policy, whether ex ante or ex post, will generate incentives to free ride. The relative magnitude of the incentives to free ride is therefore of interest. As it turns out, when redistributive effects are washed out, it is the same under both approaches. The only difference lies in the date at which those incentives manifest themselves. Fittingly, free-riding is more tempting ex post in an ex post mechanism (at the payment stage), whereas it is more tempting ex ante in an ex ante mechanism (at the agreement stage):

Proposition 3. *Assume $\lambda_t \equiv H_t'$. When tax revenues are used to compensate participants in proportion to harm—as with the liability scheme—the temptation to default on liability payments (relative to that of abandoning a Pigovian tax) is equivalent to the temptation to delay the implementation of a Pigovian tax on emissions (relative to that of delaying the implementation of the liability scheme).*

The proof of Proposition 3 can be found in Appendix A.1.

Like all international climate policies, the climate liability mechanism is a

[18] Section 5.1.1 discusses the institutions necessary to ensure an effective implementation of the liability mechanism. In turn, Section 5.2 specifically discusses the issues of commitment and credibility

voluntary one.[19] Hence, for it to be effective, there must be willing participants. This turns out to be the case under a suitably chosen liability scheme. Aggregate welfare gains do not mean that all individual countries have an incentive to participate. But it implies that *ad hoc* monetary transfers can lead to positive welfare gains for all. That being said, the mechanism we envision operates transfers based on realized harm, which may not coincide with those transfers necessary to garner participation. This is the price to pay to correct for the inequalities generated by the spatial idiosyncracies of climate harm.[20] Nevertheless, if participants are similar enough, the liability scheme results in a Pareto improvement. Moreover, the more participants, the higher the incentives to participate, which hints at the possibility of a positive unraveling effect spurred by an initial "coalition of the willing".

Proposition 4. *For any coalition of countries, there exists a liability scheme that generates aggregate benefits over the business-as-usual scenario. The larger the coalition, the closer the liability scheme can be to the efficient one, which is always beneficial when all countries participate. Moreover, the larger the coalition of participating countries, the higher the incentives to participate. In any case, if participating countries are similar enough, the gains from reduced harm always dominate redistributive transfers, so that the liability scheme results in a Pareto improvement.*

The proof of Proposition 4 can be found in Appendix A.2.

We have laid down the basic logic of a tradable climate liability scheme. The remainder of the article focuses on the 'Pigovian' liability, $\lambda_t \equiv H_t'$. It is devoted to discussing features inherent to the mechanism (Section 4) as well as issues that are not directly captured by the modeling framework (Section 5).

4 Features of the mechanism

The features of the mechanism can be separated into two categories: those related to the ex post nature of the mechanism (Section 4.1), and those related to the fact that funds are used to compensate climate harm (Section 4.2).

[19] Note that a liability scheme can be applied at the subnational level, as is the case when a federal authority has the power to impose a climate policy to its states or provinces. Then, participation is less of a concern. What we examine here is instead the situation where no superior authority exists, which is where the issue of participation is the most severe.

[20] Recall that there is no spatial relationship between climate harm and the sources of GHG emissions (e.g., Billette de Villemeur and Leroux, 2011).

4.1 Features related to the ex post nature of the approach

4.1.1 Relying on observed data yields better accuracy

The climate liability scheme exacts payments based on marginal harm and on countries' contributions to the current stock of GHGs in the atmosphere. Marginal harm is computed on the basis of past data at each period and thus inherits all of the informational advantages attached to an ex post approach. In particular, implementation does not require computing the expected sum of discounted future harm as is the case under ex ante instruments, like the Pigovian tax or an equivalent cap-and-trade scheme. With liabilities, payments associated with any given ton of GHG are instead spread over time. This allows for learning about the evolution of climate dynamics in the meantime: payments are thus based on facts rather than on hypotheses about the (possibly distant) future. We discuss possible approximations in Section 5.1.

Of course, even with observed data, disentangling between harm that is caused by human activity and harm that would have occurred naturally is itself a daunting task. The very concept of man-made harm is implicitly based on a counterfactual scenario absent anthropogenic emissions. Thus, assessing man-made harm relies on comparing an observed situation to a hypothetical one. Improvements in modelling can (with greater or lesser confidence) give us probabilistic guidance about what could have been expected to occur without anthropogenic emissions under various assumptions. Although such an endeavor may raise skepticism due to the difficulty of carrying it out accurately, note that estimating the optimal level of any ex ante instrument relies on the comparison of scenarios that are *both* hypothetical—and about the distant future, at that.

Another source of uncertainty has to do with the relationship between past emission flows and GHG concentration, as often expressed by means of a decay rate to summarize the complex geophysical and atmospheric phenomena. This decay rate depends notably on GHG concentration (Prather, 1996; IPCC, 2001). Although this relationship is most likely deterministic, an accurate computation of the evolution of atmospheric decay still requires i) an accurate prediction of these complex natural phenomena over the very long run, and ii) a complete forecast of future emissions over the very long run. The latter follows from the fact that GHG concentration at a future date will depend on today's emissions and also on all emissions until then. As a result, the future decay rate can be considered to be largely unknown to the decision maker. By contrast, an ex post policy relies only on *pas*t decay rates, which can be directly computed as the ratio of current GHG concentration (net of current emissions) to previous GHG concentrations. All calculations are thus fact based rather than scenario based.

4.1.2 Relying only on objective parameters limits misrepresentation, as does relying on a market

The implementation of the climate liability mechanism requires only knowledge of objective parameters: the stock of emissions attributable to each country and the marginal harm. Hence, there is no room for strategic manipulation by misreporting β—or any other subjective parameter—in the hopes of influencing international agreements.

Instead of countries reporting their subjective parameters formally (or revealing them through the energy spent trying to influence negotiations), the proposed liability scheme uses a market mechanism to aggregate subjective views. As a result, countries are faced with the aggregate parameters that result from market interaction, and decide to purchase or sell liabilities according to how their own stance departs from the aggregate. In essence, the market for liabilities asks countries to "put their money where their mouth is".As we discuss in Section 5.3.2, this eliminates the need for a consensus on the discount factor to be used.

Of course, this great degree of decentralization raises the question of whether countries would not tend to downplay the importance of future generations by adopting low discount factors. There are actually two questions: Can a country "adopt" a discount factor that is different from that of its population? And: Can the (actual) discount factor of a country be "too low"?

Our answer to the first question is that if decision makers display a systematic bias for the present—say, to inflate their performance while in office, for electoral concerns—the climate liability scheme would fail to implement the optimal emissions path. This is not a problem of preference elicitation, but one of political economy, which we discuss in Section 5.2.1.

The real question surrounding a country's choice of β is the second one, about whether the current generation places enough weight on the welfare of future generations. However, unless one takes a paternalistic stance—for instance, on the grounds that future generations are not represented—and decides *a priori* what the correct value of β should be, the appropriate value is actually that of the (present) population. Hence, by design, the incentives to the current generation reflect the true value of the discount factor.

4.1.3 The informational burden is displaced onto countries

The informational advantages identified above should not hide the fact that climate liabilities displace the informational burden onto countries. Indeed, even though the assessment of harm is ex post, the policy instrument is used to influence behavior ex ante. Therefore, it is up to countries to make predictions about future harm scenarios and to choose a discount factor.

Making predictions about man-made climate harm and taking a stance on the weight to be attributed to future generations are simply inescapable. Whatever the approach, climate change will remain an intertemporal problem over the very long run that is characterized by severe uncertainty. Nevertheless, we discuss in Section 5.1.1 the kind of institutional apparatus that may assist countries in making predictions.

4.1.4 Mechanism is ex post, but not retroactive (unless desired)

The key feature of a climate liability mechanism is that the prospect of being liable for future harm creates incentives to reduce emissions today. Interestingly, the mechanism does not require tracing emissions back to infinity to provide the "right" incentives, in the sense of leading to the optimal emissions pattern. It is enough, from the point of view of incentives, to account for anthropogenic emissions starting at some agreed-upon reference date only. Any starting date will work. This allows the mechanism to comply with the basic legal principle of constructive notice while providing a tool to address the normative role of historical emissions.[21]

That being said, the choice of the starting date is of importance because the large emitters of the distant past (mainly, the developed West) differ from the current—and likely future—ones. Thus, the actual distribution of climate debt will be directly affected by the chosen starting date.[22] Indeed, while incentives are unaffected, choosing an earlier starting date amounts to enacting lump-sum transfers from early emitters to more recent and future emitters. Ultimately, the climate liability scheme handles this important distributional concern in a way

[21] Obviously, for incentives to exist, "actual notice" is also needed; i.e., countries must not only be aware of the fact that they are causing harm (constructive notice), but must also be informed that they will be considered liable for future climate harm.

[22] Skeie et al (2017) confirms that the choice of a starting year is crucial in assessing countries' contribution to climate harm.

that is transparent and according to a rationale that is grounded in the historical pattern of emissions. In other words, the redistribution brought about by the choice of a starting date is strictly rooted in the climate change problem. In Section 5.3.6 we discuss the appeal of this feature for international negotiations.

4.2 Features related to the compensation of harm

4.2.1 An insurance scheme against spatial variations in harm

A distinctive feature of the mechanism is that countries receive compensation in proportion to the realized harm in each period. Thus, after compensation, each participating country must suffer only a fraction of the climate harm occurring on its own territory. As a result, the mechanism functions as an insurance scheme against idiosyncratic climate risk. It also follows that inequalities in net harm are reduced across countries.

Insurance coverage expands with the number of participants. If the systemic component of the harm function is convex, full coverage may even be achieved when participation is global.

4.2.2 Long-term risk is borne by those who are responsible for it

Under a liability scheme, long-term risk—understood as the risk of making incorrect predictions in terms of climate scenarios—is borne by those countries who are responsible for climate change. Indeed, because liability payments are based upon current marginal harm and the contribution of countries to the stock of atmospheric GHGs, it is the emitting countries that are held accountable for long-term climate risk—to the tune of their cumulative emissions—rather than those exposed to climate damage as it is the case with the ex ante approach. When liabilities are traded, the purchasing countries end up bearing this risk. However, not only will they have chosen to do so—unlike victims of climate harm[23]—but they will have gained financially through the liability price.

5 Discussions

The remaining points of discussion cannot be directly addressed with our modeling framework. After discussing implementation challenges and how they may be overcome (Section 5.1), we speculate on how governments may react to a

[23] We thank Yann Kervinio for this observation.

liability, both in terms of commitment (Section 5.2) and participation (Section 5.3).

5.1 Implementation

This section raises certain implementation aspects that must be considered in order for the mechanism to be effective. They are matters of institutional design, and strategies to avoid moral hazard regarding adaptation.

5.1.1 Institutional design

Clearly, a climate liability market requires strong institutions in order to function properly. Implementation involves five major steps: a) tracking the emissions of countries, b) issuing liabilities to emitters, c) assessing harm, d) imposing liability payments, and e) managing the transfer of liabilities within the market. Hence, the mechanism minimally requires a central authority. However, it is of note that some well-established institutions already handle several of the above-mentioned duties.

a) Reliable databases exist for historical emissions which are obtained from national accounting on the most GHG-intensive industries, like fossil fuels, cement, steel, etc. (see, e.g., International Energy Agency, 2014; CAIT Climate Data Explorer, 2015; Boden et al., 2015). From a legal perspective, implementing liability usually proves more problematic when there are many tortfeasors and possible interaction effects between their activities. However, this concern does not apply to climate liabilities. Climate science can determine the intensity of radiative forcing from the concentrations of various GHGs in the atmosphere. This knowledge, in addition to the knowledge that GHGs mix in the atmosphere, implying that the location of climate harm is independent from the source of emissions, is enough to guarantee that the emissions patterns of countries are all that are needed to establish their (relative) responsibility in climate harm. In other words, while the question of *how much climate harm* may be difficult to answer, the question of *who is causing the harm* is not.

Aspects b) and d) could be handled by the same institutional body. That organization will have to direct the flow of payments from liable countries to those who are to receive compensation. It must also be able to monitor the standing of countries with respect to their liabilities payments, in order to act as a gatekeeper to the liability market, if not to outright enforce payments. Finally, it would ensure that no country overstretches its liability position. For example, a simple but effective rule could be to impose a cap on the amount of liability a country may hold that is in proportion to that country's ability to pay. To fix ideas, note that the IMF is quite close to the institutional body just described. It is an institution that lends funds directly to national governments. As such, it routinely tracks their debt levels and handles repayment.

As for aspect e), government bonds are traded daily on international financial markets. The trade of climate liabilities would be no different.

The more challenging aspect is undoubtedly aspect c), the assessment of harm. Indeed, not only is environmental harm difficult to measure, the question of attribution—i.e., establishing how much of the harm is due to human activity—is a difficult one, that is far from being answered.[24] Nevertheless, the complexity of the climate system does not imply that we should not strive to develop a unifying framework to attribute harm to the human influence on climate.[25] Until such a framework is adopted, a pragmatic way out of the difficulty would be to rely on agreed-upon conventions about how much harm to attribute to human activity. For example, countries could agree on a given percentage, denote it x, such that $x\%$ of the harm from extreme weather events is attributed to anthropogenic climate change, and use it to calibrate liability payments. Insurance companies frequently assess damage, including when linked to climate events. Many reputable research organizations on climate science, the most prominent being the IPCC, are focused on assessing and forecasting climate harm, including for non-extreme weather events like the gradual regional changes in temperature (slow-onset events).[26] One may be concerned that such an institutional body may be subject to the influence of certain governments. However, note that the IPCC already makes climate predictions and estimations of the remaining "carbon budget" that have high-stake consequences for many countries, all the while retaining a reputation for being independent.

[24] Observe first that, as is well-known in cost-benefit analysis, there are several notions of "value" (market value, use value, option value, existence value, etc.). Hence, evaluating harm requires at least to precisely define the appropriate notion(s) of value to be considered. Doing so is beyond the scope of this work. However, our approach can accommodate any notion of value. Yet, this may require some careful handling, as with the extinction of species, say polar bears, which have both a direct use value (for arctic hunter populations), an existence value (for most of the planet's population), and even an intrinsic value. This will result in differentiated harm across populations.

[25] Allen et al. (2007) points to such a possibility. Allen (2003) had previously been one of the first to call attention to the magnitude of the challenge.

[26] For such slow-onset events, it may also be useful to resort to educated rules of thumb initially. For example, one could approximate marginal harm by dividing the trend in the past recent years (so as to average out the randomness of the harm) by the increase in the atmospheric stock of GHGs over the same time period.

Also, an organization like the IPCC could provide regular forecasts upon which countries would base their own decisions, thus saving the costs associated with having these predictions be made redundantly by each country.

5.1.2 Moral hazard and incentives for adaptation

Because liability revenues are used to (partially) compensate climate harm, it presumably undermines the incentives of countries to take adaptive measures to reduce climate impacts within their borders. It is the very same moral hazard problem that plagues insurance markets, for instance, which explains why insurance policies rarely allow for complete coverage. Similarly, imposing a threshold below which climate harm is *not* compensated—i.e., a deductible of sorts—would restore the incentives to adapt at the margin. It would also avoid costly litigation (over the magnitude of harm) for relatively small events for which the link with climate change is difficult to assess. More generally, limiting compensation to sizable harms would obviously ease the measurement problem.

Note that the moral hazard problem does not prevent insurance markets from being both profitable and beneficial to society. Despite the downward incentives on adaptation, we still believe that it is desirable to compensate countries affected by man-made climate harm, especially because those who are most impacted are often not those who emit the most. In essence, we believe the public bad nature of the climate change problem supersedes this moral hazard concern.

Moreover, just as there can be shared responsibility in an accident, shared responsibility rules could be introduced to limit compensation in the absence of adaptation, which would restore the incentives to adapt. For example, in cases of obvious negligence, like knowingly developing flood-prone land, countries could be denied compensation for the associated harm.

On top of these considerations, it is worth noting that holding climate liabilities actually provides incentives *for* adaptation. It effectively transfers incentives from the countries that incur the climate harm to the liability bearers. The existence of a global liability price will generate incentives for countries holding high liability positions to invest in adaptation wherever it is most productive (including outside of its borders), thus fostering an efficient allocation of adaptation efforts.[27] These incentives for adaptation will ultimately translate into lower expected harm.

[27] The argument is similar in spirit to Harstad (2012).

5.2 Commitment and credibility

Proposition 3 establishes that the relative temptation to renege on the climate liability mechanism (by defaulting on liability payments) is equivalent to the relative temptation to postpone the Pigovian carbon tax, at least theoretically. This section discusses more broadly the vulnerability of the climate liability mechanism to the issue of commitment and the related issue of credibility.

5.2.1 Unconcern for future generations and incentives for renegotiation

An obvious limitation of the liability approach is that it requires countries to commit to honoring payments over the very long run. This raises two political economy-related problems, one across generations and one across countries.

Across generations: Political short-termism First, the relative short-sightedness of governments, due to relatively short electoral cycles, may induce them to neglect the burden they impose upon their successors. As a result, governments may tend to underestimate the (future) cost of emitting (today) and be inclined to allow more emissions than would be in the country's (intertemporal) interest. One can expect this lack of concern for future generations to undermine the effectiveness of a climate liability mechanism: Governments would emit more GHGs than is globally optimal, as if they were subject to a much weaker—or nonexistent—climate policy.

Unfortunately, no climate policy is immune to the shortsightedness of politicians. This is simply due to intertemporal nature of the climate problem and the long timespans involved. For instance, in the case of ex ante policy instruments, it simply takes on different forms. For fear of losing political support from businesses, politicians may be reluctant to implement an ex ante policy that has real bite. This translates into imposing a price on emissions, as measured by effective tax rates and cap-and-trade prices, that is widely considered by experts to be far below most reasonable estimates of the social cost of carbon (e.g., Pindyck, 2016; Ricke et al., 2018), resulting in a weaker policy. Alternatively, the policy may be postponed altogether, resulting in a nonexistent policy.

As already mentioned, by making it visible that debt that accumulates, we speculate that voters will be more sensitive to the climate issue, thus disciplining politicians in turn.

Across countries: Temptation to default on liability payments The second problem arises between the liable countries and those who are to be compensated. As time goes on and as the financial burden increases over time for liable countries, one may expect them to choose to default on their liability payments. This could leave the countries affected by climate change without the compensations they were promised, with possibly dramatic consequences to their populations.

This is a legitimate concern that should not be ignored. Indeed, Bulow and Rogoff (1989) assert that sovereign debt cannot be supported by a country's reputation for repayments when default carries only a ban from ever borrowing in financial markets. The argument is that, by defaulting, a country can save upon repayments, increase current consumption, and replicate the same consumption pattern for the future without borrowing—when markets are complete, by making use of securities.

In the presence of uninsurable risks, however, a country may prefer to repay its debt so as to maintain access to the opportunities of risk diversification that the asset market provides. As several authors pointed out (Eaton, 1990, Chari and Kehoe, 1993), the argument of Bulow and Rogoff hinges on the fact that the defaulting country can re-lend its outstanding debt to yet another agent in exchange for a guaranteed stream of income. In fact, once one recognizes that this secondary debtor is not itself committed, Hellwig and Lorenzoni (2009) have shown that debt can actually be sustainable, which is more likely when interest rates are low.

Very recently, Bloise, Polemarchakis and Vailakis (2017) extended the result of Bulow and Rogoff to the incomplete market case. Clearly, the debate on the effectiveness of borrowing bans to discipline debtors is far from being closed. A literature review by Panizza, Sturzenegger and Zettelmeyer (2009) concludes that the problem of sovereign debt does not so much arise because of the lack of external sanctions, but is largely determined the domestic costs of default. This indirectly supports accounting explicitly for the costs of debt, as we have done in the model.

Under a climate liability scheme, failure to honor liability payments is no different than defaulting on financial debt. This is an important feature of the mechanism over other economic instruments: a climate liability market actually provides incentives to honor liability payments through the "financialization" of climate harm. Defaulters would not simply have to bear the moral consequences of reneging on their commitment, but would also suffer financial setbacks. Furthermore, we believe that the temptation to default in a liability scheme is tempered by the fact that defaulting makes a country ineligible for compensations that insure, at least partially, against the risk of climate harm.

Interaction between incentives and participation/commitment. The political-economy problems just discussed unfold along two different dimensions. On the one hand, the intergenerational problem yields low incentives to reduce emissions, under both a carbon tax and a climate liability market. On the other hand, the cross-country problem yields a participation/commitment dilemma: a participation problem for the carbon tax (temptation to delay indefinitely) and a commitment problem for the climate liability market (temptation to default on payments), respectively.

However, both aspects are interwoven. A low carbon tax, while providing limited incentives to reduce emissions, actually makes participation easier. Conversely, delayed participation would make any given carbon tax rate even lower relative to the optimal one.[28] Summing up, with a carbon tax, the incentives problem alleviates the participation problem whereas the latter reinforces the former.

The incentives and commitment problems are also related for a climate liability market, although somewhat differently. A lack of concern for future generations yields too much GHGs to be emitted, which exacerbates the commitment problem (due to holding more liability). Conversely, anticipating default on liability payments, countries have then little incentive to reduce emissions today. In other words, contrasting with the carbon tax, the incentives and the commitment problems reinforce each other. In our opinion, this is the most significant limitation of the climate liability market.

5.2.2 The mechanism requires making provisions for the future

A consequence of the financial nature of the debt and the associated increased government accountability (Section 5.2.1) is that governments will be compelled to save resources for future generations. To be clear, this does not imply that governments should immobilize capital in the event of future climate harm, which would be an inefficient use of resources. As we discuss in Section 5.3.4, countries are free to choose how to levy the monies they deem necessary to face their future liability payments. Moreover, governments can choose to invest

[28]The fact that the social cost of carbon increases over time is supported by most integrated assessment models (see, e.g., US Government, 2013)

these funds productively. What matters for a country, and in particular for its credit rating, is its overall balance sheet (the value of its assets net of its debt, including climate liabilities).

One may question the ability of governments to make provisions for the future, in light of the fact that several countries already face difficulties in honoring their promises to retirees, for instance. However, it is noteworthy that this tends to be the case of countries that have adopted a pay-as-you-go retirement scheme. In such systems, the current working generation bears the charge of the benefits to the previous one. There is no explicit debt. Hence, the financial balance is not monitored and, therefore, there are few options to cope with the imbalance created by the aging of populations.

By contrast, in a fully-funded pension scheme, the balance between contributions and future benefits is closely monitored. Pension funds are important financial actors, the performance of which is the object of constant scrutiny. Contributions are regularly adjusted in a system of defined benefits, whereas benefits follow market returns in a system of defined contribution. Either way, there are no lasting imbalances.

We argue that the very act of making explicit (and financial) the otherwise moral debt to future generations results in a closer examination of wealth transfers across generations. Keeping with the analogy, a carbon tax would be tantamount to a 'pay-as-you-emit' scheme, whereas a climate liability market resembles a fully-funded system, where liability returns are studied closely by financial markets.

5.2.3 Safeguards are needed to prevent opportunistic behavior

One may be concerned that because liability payments are deferred to the future it will encourage some countries to participate in the scheme only to quit before large payments are required. The worst type of opportunistic behavior would be a 'take-the-money-and-run' tactic: massively purchase liability (and be paid to endorse this extra liability) and exit immediately, thus keeping the money while not honoring the debt. This is a serious concern, that must be addressed for the mechanism to perform as intended.

First of all, recall that the mechanism restricts access to the liability market. Access is limited to participating countries and, among them, to countries that are honoring their liability payments. Any participating country that is behind on its liability payments will be excluded from accumulating more liabilities through the market until it has fully honored its liability payments.

Secondly, observe that, by exiting, a country forgoes any claims to the compensation afforded by the liability scheme. By itself, this feature of the mechanism does not prevent take-the-money-and-run tactics. However, it does shed light on *which* countries may be tempted to behave as such: countries that are net contributors to the system. They are also the countries that are least likely to participate in the first place.

Finally, one could impose additional safeguards to prevent opportunistic behavior. For example, one could limit the amount of liabilities a country can purchase, by imposing a threshold net-liability-to-GDP ratio, common to all, below which a country would have to stand in order to be allowed to purchase more liabilities.

5.3 Participation

Proposition 4 established that a market for tradable climate liabilities theoretically allowed for Pareto improvements over the business-as-usual scenario. This section takes a broader view and speculates about aspects that may affect the likelihood of countries to participate in the scheme.

5.3.1 Payments are spread over time

The mechanism requires liability bearers to pay over time in proportion to their stock of GHG in the atmosphere. By spreading payments, the initial phase of a climate agreement is a smooth process rather than requiring emitters to pay "up front" for the expected future harm attached to their emissions. This is likely to ease countries into participating.

5.3.2 The mechanism decentralizes important parameters

An important advantage of the liability scheme is that payments do not depend upon forecasts of man-made climate harm nor on any discount factor. Accordingly, countries do not need to agree along these two dimensions in order to participate in the climate liability market. Each country can express its own beliefs about future climate harm and its own attitude towards future generations by choosing the amount of liability it wishes to hold.[29]

[29]Shavell (1992) argues that, left to their own devices to gather information, potential tortfeasors may lack efficient incentives to obtain information about risk under a negligence rule, where duty of care must ascertained. However, this drawback does not apply to strict liabil-

This can have strong implications in terms of international participation because it circumvents the controversies over the expected impact of climate change and the unending debate over the appropriate discount factor.[30] By reducing the number of dimensions on which countries need to agree, one can expect a climate liability market to greatly facilitate the negotiation process toward a powerful climate treaty.

5.3.3 The mechanism gives financial incentives to participate in the market

A climate liability market is likely to promote participation from countries that may initially have little interest in a scheme aimed at reducing GHG emissions. This follows from the profit opportunities that come with participating in any market. In fact, to take an extreme example, even a country that (sincerely) denies the existence of the climate problem will want to participate in a climate liability market: to this country, the purchase of any unit of liability comes with a windfall of p, for a much lower (perceived) cost. Either for genuine environmental concern or for financial gain, a well-designed climate liability market can be expected to garner the participation of many countries.

By contrast, any mechanism that requires countries to pay ex ante—i.e., before harm is realized—is a political nonstarter for climate skeptics.

5.3.4 States remain sovereign in how to fund their liability payments

A difficulty attached to global endeavors is that governments do not like to be bound in their policy choices. As it turns out, a climate liability approach ultimately gives countries full freedom in how to tackle the climate issue. This stems from the fact that climate liabilities effectively do nothing more than to convert emissions into financial debt. It is then up to each country to devise its own policy in order to achieve the desired level of emissions and to finance the corresponding debt. This can be done through a (national) tax on emissions, auctioning tradable allowances, or any other instrument (or mixture of instruments) of that country's choice.

---

ity, which is the case we consider. To the contrary, strict liability holds potential tortfeasors responsible till the end, regardless of care. Rather, it is taxation which removes responsibility—and, thus, incentives to gather information—entirely.

[30] See, e.g., Weisbach and Sunstein, 2008, for a review of the terms of the debate.

5.3.5 Liability payments are uncertain

Countries may be reluctant to participate in a liability mechanism on the grounds that countries are not contracting on monetary amounts determined ex ante, but on payments that may vary over time, because they depend on future marginal climate harm. This uncertainty may have negative consequences on economies, due to uncertainty possibly scaring away investors.

This is the price to pay to be able to charge marginal harm using the most accurate information possible. Any ex ante mechanism will charge a price that is known at the time of emission, but this apparent advantage hides the fact that climate harm is still a phenomenon we are learning about. With an ex ante mechanism, the difference between actual marginal harm and expected marginal harm at the time of emission is effectively borne by the emitters themselves—should we discover that climate change is not as severe a problem as initially thought—or, worse, by those countries who bear the brunt of the consequences—should climate change turn out to be an even worse problem than anticipated. By contrast, with a liability scheme, emitters pay the actual marginal harm regardless of prediction errors.

The liability scheme translates into financial terms part of the (climate) risk that countries face under an ex ante mechanism. In other words, it does not *introduce* risk; rather, it makes explicit risk that was already there. The difference is that emitting countries bear part of the risk with a liability mechanism (as mentioned in Section 4.2.2) whereas it is those countries that suffer the harm that bear all of the risk under an ex ante mechanism.

5.3.6 The choice of a starting date assigns responsibility for historical emissions in a diplomatic way

A climate liability scheme may actually ease international negotiations. Indeed, it allows countries to address issues of fairness without impinging on efficiency. By contrast, negotiations revolving around ex ante instruments—say, a carbon tax or, more generally, on emission targets—necessarily trade off fairness for efficiency, which leads to compromises that are deemed unacceptable by some on efficiency grounds and by others for fairness reasons. The coalition structure generated by the Kyoto Protocol, with developed countries on one side and developing countries on the other, is telling.

From a political point of view, an advantage of the climate liability scheme is that it brings to light a stumbling block of climate negotiations that, although critical, is usually not explicitly addressed. Indeed, it offers a built-in framework to decide the issue of responsibility for past emissions. This allows to depart from the crude solutions usually put together to face the obvious heterogeneity of countries in development status and wealth, like partitioning countries into different groups or "Annexes". By setting a date from which emissions will be accounted for, a climate liability mechanism explicitly acknowledges historical responsibility and specifies the degree to which each country should be held accountable in relation to the amount of past emissions. The earlier this date, the higher the burden to more developed countries, as per the high correlation of historical emissions with current GDP levels. Moreover, fine-tuning the starting date allows for smooth adjustments in the responsibility borne by countries. Hence, responsibility is no longer binary.

Finally, from a qualitative standpoint, the choice of a starting date differs significantly from choosing thresholds to partition countries into groups. First, the choice of a starting date affects all countries, and not just the countries at the margin of belonging or not to a category (i.e., located close to a threshold). Next, adjustments in the starting date affect countries gradually, as opposed to having decisive consequences to those same marginal countries while others are not affected at all. In terms of climate agreements, negotiations about the starting date re-balances the stakes and spreads them from a few countries to all nations instead of entrenching the positions of opposed coalitions.

6 Conclusion

We have proposed a new approach to climate policy based on climate liabilities between countries and have demonstrated that it can lead to the same emissions pattern as a Pigovian ex ante instrument. In fact, although it adopts an ex post approach, unlike traditional economic instruments, we claim that the mechanism is merely an unconventional interpretation of the familiar Pigovian principle. Recall that Pigovian taxation was historically developed in a static setting, where the externality could plainly be observed. The climate problem adds two dimensions to the externality: duration and uncertainty. The Pigovian logic can therefore be revisited accordingly. If we take the Pigovian logic to mean the internalization by agents of all aspects related to the externality, then our proposal applies this logic in full by relying on the countries' own views about discounting and on their own beliefs about future damage to induce them to internalize the externality.

By turning environmental debt into financial debt, the responsibility of countries becomes both explicit and more difficult to ignore, thus increasing government accountability. It becomes explicit because the financial value of the debt is indeed a market assessment of future harm. It becomes difficult to ignore because “the market is watching”. The market (liability) price aggregates all available private information and makes it public. Thus, the total value of liabilities at any current date is the best possible estimate of the cost of future climate harm.

We are able to make some qualitative predictions about participation (Proposition 4). Quantitative predictions about the size of the welfare gains over the business-as-usual scenario would require making use of simulations. In this vein, Antimiani et al. (2017) develop a CGE model to assess whether the Green Climate Fund could work as an effective transfer mechanism in securing global participation. There, the usage of the Fund is split between compensating abatement (i.e., economic loss from reduced production) and funding adaptation (to reduce climate harm). By contrast, the transfer scheme we propose focuses solely on compensating harm.

We have considered climate liabilities as an alternative to other, ex ante economic instruments like a carbon tax. However, in practice, the combination of the two is generally beneficial. In a paper on liability and taxation, Shavell (2011) argues that “liability should be employed fully, with taxation taking up the slack”. In our setting, it is quite clear that any weighted average of the 'Pigovian' liability and taxation schemes would also implement the desired level of emissions. Therefore, without sacrificing efficiency, one can fine-tune the relative weights of both instruments to achieve the most desirable mix of advantages and drawbacks brought about by each.

References

[1] Allen, M. (2003) “Liability for climate change.” *Nature*, 421(6926): 891-892.

[2] Allen, M.R., Pall, P., Stone, D.A., Scott, P., Frame, D.J., Min, S-K., Nozawa, T. and S. Yukimoto (2007) “Scientific challenges in the attribution of harm to human influence on climate.” *Pennsylvania Law Review*, **155**(6): 1353-1400.

[3] Ambec, S. and L. Ehlers (2016) “Regulation via the Polluter Pays Principle”, *The Economic Journal*, **126**, 884-906.

[4] Antimiani, A., V. Costantini, A. Markandy, E. Paglialung, and G. Sforna (2017), “The Green Climate Fund as an effective compensatory mechanism in global climate negotiations”, *Environmental Science and Policy*, **77**, 49-68.

[5] Benchekroun, H. and N. van Long (1998) “Efficiency inducing taxation for polluting oligopolists”, *Journal of Public Economics*, **70**, 325-342.

[6] Billette de Villemeur, E. and J. Leroux (2011) “Sharing the Cost of Global Warming”, *The Scandinavian Journal of Economics*, **113**(4), 758-783.

[7] Billette de Villemeur, E. and J. Leroux (2012) “CO_2: Tax Now, Pay Later!”, *Tax Notes International*, **66**, 49-53.

[8] Bloise, G., Polemarchakis, H., & Vailakis, Y. (2017). “Sovereign debt and incentives to default with uninsurable risks”, *Theoretical Economics*, **12**(3), 1121-1154.

[9] Bobtcheff, C., Chaney, T., & Gollier, C. (2016), “Analysis of Systemic Risk in the Insurance Industry”, *The Geneva Risk and Insurance Review*, 41(1), 73-106.

[10] Boden, T.A., G. Marland, and R.J. Andres. 2015. *Global, Regional, and National Fossil-Fuel CO2 Emissions.* Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. doi 10.3334/CDIAC/00001_V2015

[11] Bovenberg, A.L. and B.J. Heijdra (1998) “Environmental tax policy and intergenerational distribution.” *Journal of Public Economics*, **67**, 1–24.

[12] Buchholz, W. and Schumacher, J. (2008) “Discounting the Long-distant Future: A Simple Explanation for the Weitzman-Gollier Puzzle”. CESifo Working Paper 2357

[13] Bulow, J., & Rogoff, K. (1989). “A constant recontracting model of sovereign debt”, *Journal of political Economy*, **97**(1), 155-178.

[14] Cairns, R. (2004). “Green Accounting for an Externality: Pollution at a Mine.” *Environmental and Resource Economics* **27**: 409-427.

[15] Cairns, R. and P. Lasserre (2006) "Implementing Carbon Credits for Forests based on Green Accounting", *Ecological Economics* **56**, 610-621.

[16] CAIT Climate Data Explorer. 2015. Washington, DC: World Resources Institute. Available online at: http://cait.wri.org.

[17] Calabresi, G., *The Cost of Accidents: A Legal and Economic Analysis*, New Haven, CN.: Yale University Press, 1970.

[18] CDP (2017) "The Carbon Majors Database - CDP Carbon Majors Report 2017" https://6fefcbb86e61af1b2fc4-c70d8ead6ced550b4d987d7c03fcdd1d.ssl.cf3.rackcdn.com/cms/reports/documents/000/002/327/original/C Majors-Report-2017.pdf?1501833772 (Accessed Sept. 25th 2018)

[19] Chari, V. V., & Kehoe, P. J. (1993). "Sustainable plans and mutual default", *The Review of Economic Studies*, **60**(1), 175-195.

[20] Dennig, F., Budolfson, M. B., Fleurbaey, M., Siebert, A., & Socolow, R. H. (2015). Inequality, climate impacts on the future poor, and carbon prices. Proceedings of the National Academy of Sciences, 112(52), 15827-15832.

[21] Diamond, D. W., & Dybvig, P. H. (1983) "Bank runs, deposit insurance, and liquidity", *Journal of political economy*, 91(3), 401-419.

[22] Eaton, J. (1990). "Sovereign debt, reputation, and credit terms", *National Bureau of Economic Research Working Paper,* No. 3424.

[23] Engen, E.M. and R.G. Hubbard (2004) "Federal Government Debt and Interest Rates," in *NBER Macroeconomics Annual 2004, Vol. 19*, Mark Gertler and Kenneth Rogoff Eds., MIT Press.

[24] Farhi, E., & Tirole, J. (2012), "Collective moral hazard, maturity mismatch, and systemic bailouts", *American Economic Review*, 102(1), 60-93.

[25] Faure, M. and M. Peeters Eds. (2011), *Climate Change Liability*, in *New Horizons in Environmental and Energy Law* series, Edward Elgar Publishing.

[26] Frank, W. (2017) "The Huaraz Case (LLuiya v. RWE) - German court opens recourse to climate law suit against big CO2-emitter" http://blogs.law.columbia.edu/climatechange/2017/12/07/the-huaraz-case-lluiya-v-rwe-german-court-opens-recourse-to-climate-law-suit-against-big-co2-emitter/#_ftnref1 (Accessed Sept. 25th 2018)

[27] Gampfer, R. (2014) "Do individuals care about fairness in burden sharing for climate change mitigation? Evidence from a lab experiment." *Climatic Change* **124(1)**, 65-77.

[28] Gampfer, R., E. Gsottbauer and A. Delas (2014) "Interactive experiment: compensation for loss and damage as a means to shore up climate ambition?" mimeo ETH-Zurich.

[29] Gerlagh, R. and M. A. Keyzer (2001) "Sustainability and the intergenerational distribution of natural resource entitlements," *Journal of Public Economics*, **79**, 315-341.

[30] Gersbach, H., and R. Winkler (2011) "International emission permit markets with refunding," *European Economic Review*, **55** 759-773.

[31] Gersbach, H., and R. Winkler (2012) "Global Refunding and Climate Change," *Journal of Economic Dynamics and Control*, **36** 1775-1795.

[32] Gollier, C. and R. Zeckhauser (2005) "Aggregation of Heterogeneous Time Preferences", *Journal of Political Economy*, **113**, 878-896.

[33] Gollier, C., and M.L. Weitzman (2010) "How should the distant future be discounted when discount rates are uncertain?" *Economics Letters*, **107**(3), 350-353.

[34] Harstad, B. (2012) "Buy Coal! A Case for Supply-Side Environmental Policy," *Journal of Political Economy*, **120**, 77-115.

[35] Heal, G. and A. Millner (2014) "Agreeing to disagree on climate policy", *PNAS*, **111**(10), 3695-3698.

[36] Heal, G. and A. Millner (2018) "Discounting by committee", *Journal of Public Economics*, **167**, 91-104.

[37] Hellwig, C., & Lorenzoni, G. (2009) "Bubbles and self-enforcing debt", *Econometrica*, **77**(4), 1137-1164.

[38] Hiriart, Y., Martimort, D. and J. Pouyet (2004) "On the optimal use of ex ante regulation and ex post liability," *Economics Letters*, **84**, 231-235.

[39] Hiriart, Y., Martimort, D. and J. Pouyet (2010) "The public management of risk: Separating ex ante and ex post monitors," *Journal of Public Economics*, **94**, 1008-1019.

[40] International Energy Agency (IEA). 2014. *CO2 Emissions from Fuel Combustion* (2014 edition). Paris, France: OECD/IEA. Available online at: http://data.iea.org/ieastore/statslisting.asp. © OECD/IEA, [2014].

[41] IPCC (2001), Third Assessment Report, *Working Group I: The Scientific Basis*, Chap. 4. http://www.ipcc.ch/ipccreports/tar/wg1/index.php?idp=127 (accessed Dec. 7th 2016)

[42] Jackson, M. and L. Yariv (2015) "Collective Dynamic Choice: The Necessity of Time Inconsistency", *Amer. Econ. J.: Microeconomics*, **7**, 150-178.

[43] Kessler, D. (2014) "Why (re) insurance is not systemic", *Journal of Risk and Insurance*, **81**(3), 477-488.

[44] Koopmans, T. C. (1960) "Stationary ordinal utility and impatience", *Econometrica*, **28**, 287–309.

[45] Kolstad, C.D., T.S. Ulen and G.V. Johnson (1990), "*Ex Post* Liability for Harm vs. *Ex Ante* Safety Regulation: Substitutes or Complements?", *American Economic Review*, **80(4)**, 888-901.

[46] Laubach, T. (2009) "New Evidence on the Interest Rate Effects of Budget Deficits and Debt," *Journal of the European Economic Association*, **7,** 858-885.

[47] Lord, R., S. Goldberg, L. Rajamani and J. Brunée Eds (2012), *Climate Change Liability: Transnational Law and Practice*, Cambridge University Press.

[48] Magill, M., Quinzii, M. and J.-C. Rochet (2015), "A Theory of the Stakeholder Corporation," *Econometrica*, **83(5)**,1685-1725.

[49] Moulin, H. (2002) "Axiomatic cost and surplus sharing," in: K. J. Arrow & A. K. Sen & K. Suzumura (ed.), Handbook of Social Choice and Welfare, edition 1, volume 1, chapter 6, pages 289-357 Elsevier.

[50] Nordhaus, W. (2007) "A Review of the *Stern Review on the Economics of Climate Change*", *Journal of Economic Literature*, vol. **XLV**, 686-702.

[51] Panizza, U., Sturzenegger, F., & Zettelmeyer, J. (2009), "The economics and law of sovereign debt and default", *Journal of Economic Literature*, **47**(3), 651-98.

[52] Pindyck, R. S. (2016) "The Social Cost of Carbon Revisited," National Bureau of Economic Research, no. 22807.

[53] Prather, M.J. (1996) "Time scales in atmospheric chemistry: Theory, GWPs for CH4 and CO, and runaway growth," *Geophysical Research Letters*, **23**(19).

[54] Ricke, K., L. Drouet, K. Caldeira and M. Tavoni (2018) "Country-level social cost of carbon," *Nature Climate Change*, 24 sept. 2018.

[55] Skeie, R.B., J. Fuglestvedt, T. Berntsen, G.P. Peters, R. Andrew, M. Allen and S. Kallbekken (2017) "Perspective has a strong effect on the calculation of historical contributions to global warming", *Environmental Research Letters*, **12**(2).

[56] Shavell, S. (1992), "Liability and the Incentive to Obtain Information about Risk", *The Journal of Legal Studies*, **21**(2), 259-270.

[57] Shavell, S. (2011), "Corrective taxation versus liability", *American Economic Review Papers and Proceedings*, **101**(3), 273-276.

[58] Shleifer, A., *The Failure of Judges and the Rise of Regulators*. MIT Press, 352 pages, 2012.

[59] Swiss RE Institute, *Sigma 2/2017: Natural Catastrophes and Man-Made Disasters in 2016: A Year of Widespread Damages*, March 2017.

[60] UNFCCC (2015) *Paris Agreement* https://unfccc.int/files/meetings/paris_nov_2015/application/pdf/paris Accessed on Dec. 11th 2018

[61] UN Statistical Commission (2016) "Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators," unstats.un.org/unsd/statcom/47th-session/.../2016-2-IAEG-SDGs-Rev1-E.pdf (accessed Dec. 7th 2016)

[62] *University of Pennsylvania Law Review*, Symposium: Responses to Global Warming: The Law, Economics and Science of Climate Change, **155**(6), 2007.

[63] U.S. Goverment, Interagency Working Group on Social Cost of Carbon (2013) "Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis"

http://www.whitehouse.gov/sites/default/files/omb/assets/inforeg/technical-update-social-cost-of-carbon-for-regulator-impact-analysis.pdf (accessed Dec. 7th 2016)

[64] von Below, D., F. Dennig, and N. Jaakkola (2016) "The climate-pension deal: An intergenerational bargain," mimeo, University of Nottingham. https://www.nottingham.ac.uk/climateethicseconomics/documents/papers-workshop-2/jaakkola-et-al.pdf

[65] Wachtel, P. and J. Young (1987) "Deficit Annoucements and Interest Rates," *The American Economic Review*, **77**(5), 1007-1012.

[66] Weisbach, D. and C.R. Sunstein (2008) "Climate Change and Discounting the Future: A Guide for the Perplexed," *Yale Law & Policy Review*, **27(2)**, Article 6. Available at: http://digitalcommons.law.yale.edu/ylpr/vol27/iss2/6 (accessed Dec 7th 2016)

[67] Weitzman, M.L. (1976), "On the Welfare Significance of National Product in a Dynamic Economy", *Quarterly Journal of Economics* **90**, 156-162.

[68] Weitzman. M.L. (1998) "Why the Far-Distant Future Should be Discounted at its Lowest Possible Rate", *Journal of Environmental Economics and Management*, **36**, 201-208.

[69] Weitzman M.L. (2001) "Gamma Discounting", *American Economic Review*, **91**, 260-271.

[70] Weitzman, M.L. (2017), "On a World Climate Assembly and the Social Cost of Carbon", Harvard University, mimeo, https://scholar.harvard.edu/weitzman/publications/world-climate-assembly (accessed July 6th 2017)

[71] Zuber, S. (2011) "The aggregation of preferences: can we ignore the past?" *Theory and Decision*, **70(3)**, 367-384.

A Appendix

This appendix provides the proofs of Proposition 3 and Proposition 4.

A.1 Proof of Proposition 3

The public good character attached to climate policy means that all climate instruments are subject to the problem of free riding. In other words, at the individual level, one is always tempted to abandon (or to not implement) a climate policy, regardless of the instrument (liability or tax). This section compares in turn the relative temptation to abandon each mechanism and the relative temptation to delay the implementation of each mechanisms. For the sake of isolating the issue, we shall assume that countries are homogeneous in expectations and share a common discount factor, β. For simplicity of exposition, we also assume the initial stock to be zero, so that: $S_t^i = \sum_{s=0}^{t} \gamma^{t-s} e_s^i$ for all i. The liability instrument we shall consider is the 'Pigovian' one, where emission stocks are priced at current marginal harm: $\lambda_t \equiv H_t'$. This appendix makes heavy use of the fact that this liability scheme implements the same emissions pattern as the Pigovian tax.

A.1.1 On the relative temptation to abandon the policy

On the one hand, a drawback of the liability scheme is that countries face an increasing temptation to default on their accumulated climate debt. On the other hand, because climate harm is likely to increase over time, there is also an increasing temptation to not enact a Pigovian tax.

Denote by M the set of countries implementing the policy, either the liability scheme or a Pigovian tax. The expected cost of defaulting on climate debt from date T onward—or, alternatively, the benefits to continued participation—is equal to:[31]

$$\Delta^{M,i}_{liability,T} = E_T \left\{ \sum_{t=T}^{+\infty} \beta^{t-T} \left[b_t^i \left(e_t^{M,i} \right) - \lambda_t S_t^{M,i} + \lambda_t \left(\sum_{j \in M} S_t^{M,j} \right) \frac{h_t^i \left(S_t^M \right)}{\sum_{j \in M} h_t^j \left(S_t^M \right)} - h_t^i \left(S_t^M \right) \right] \right. \tag{12}$$

$$\left. - \sum_{t=T}^{+\infty} \beta^{t-T} \left[b_t^i \left(e_{t,T}^{M \setminus i,i} \right) - h_t^i \left(S_{t,T}^{M \setminus i} \right) \right] \right\}. \tag{13}$$

where we abuse notations slightly and denote by $e_t^{M,i}$ and $S_t^{M,i}$ the carbon flow and stock of Country i when countries in M implement the liability scheme (which are the same as when countries in M implement the Pigovian tax). Also,

[31] Recall that defaulting on climate debt makes one ineligible for compensatory transfers.

$S_t^M = \sum_{j\in N} S_t^{M,j}$ is the total stock in the atmosphere at time t. Finally, $e_{t,T}^{M\setminus i,i}$ refers to the carbon flow of Country i and $S_{t,T}^{M\setminus i}$ to the total stock, respectively, after Country i has exited the mechanism at date T.

We now turn to the Pigovian tax. To disentangle the temporal aspects from the redistributive aspects, we shall assume that all tax revenues are devoted to compensating participating countries in proportion to harm, exactly as under the liability scheme. This is feasible due to the fact that the Pigovian tax is exactly equal to the sum of all future expected discounted marginal harm, assuming perfect foresight and a rate of return r such that $\beta = 1/(1+r)$. More precisely, and using the fact that $\lambda_t \equiv H_t'$, the financial value at date T of tax revenues made until then is equal to:

$$\sum_{t=0}^{T} \beta^{t-T}\tau_t e_t^{M,i} = \sum_{t=0}^{T} \beta^{t-T} E_t \left\{ \sum_{u=t}^{+\infty} \beta^{u-t}\gamma^{u-t}\lambda_u \right\} e_t^{M,i} \tag{14}$$

$$= \sum_{t=0}^{T} \beta^{-T} \sum_{u=t}^{+\infty} E_t \left\{ \beta^u \gamma^{u-t}\lambda_u \right\} e_t^{M,i} \tag{15}$$

$$= \sum_{u=0}^{T} \beta^{-T} \sum_{t=0}^{u} E_u \left\{ \beta^u \gamma^{u-t}\lambda_u \right\} e_t^{M,i} + \sum_{u=T+1}^{+\infty} \beta^{-T} \sum_{t=0}^{T} E_u \left\{ \beta^u \gamma^{u-T+T-t}\lambda_u \right\} e_t^{M,i} \tag{16}$$

$$= \sum_{u=0}^{T} \beta^{u-T} E_u \left\{ \lambda_u \right\} S_u^{M,i} + \sum_{u=T+1}^{+\infty} \beta^{u-T} E_u \left\{ \lambda_u \gamma^{u-T} \right\} S_T^{M,i} \tag{17}$$

$$> \sum_{u=0}^{T} \beta^{u-T} \lambda_u S_u^{M,i} \tag{18}$$

where (16) obtains using a permutation on the domain of summation and makes use of the perfect foresight assumption. Expression (18) is the financial value, at date T, of all liability payments. The gap between (14) and (18), $\sum_{u=T+1}^{+\infty} \beta^{u-T}\gamma^{u-T}\lambda_u S_T^{M,i}$, is the financial value of the extra monetary outlays imposed up front by the Pigovian tax relative to the liability scheme. It is also the value of tax revenue that should be saved by countries in order to be able to compensate future harm.

Because we assume that the countries are compensated in the same manner as with the liability scheme, the expected cost of abandoning the tax—or, again, the benefits to continued participation—at a given date T for a country $i \in M$

is equal to:

$$\Delta_{tax,T}^{M,i} = E_T \left\{ \sum_{t=T}^{+\infty} \beta^{t-T} \left[b_t^i \left(e_t^{M,i} \right) - \tau_t e_t^{M,i} + \lambda_t \left(\sum_{j \in M} S_t^{M,j} \right) \frac{h_t^i \left(S_t^M \right)}{\sum_{j \in M} h_t^j \left(S_t^M \right)} - h_t^i \left(S_t^M \right) \right] \right. \tag{19}$$

$$\left. - \sum_{t=T}^{+\infty} \beta^{t-T} \left[b_t^i \left(X_{t,T}^{M \setminus i,i} \right) - h_t^i \left(S_{t,T}^{M \setminus i} \right) \right] \right\}. \tag{20}$$

The rest of the argument makes use of the fact that $e_t^{M,i}$ and $S_t^{M,i}$ (resp. $e_t^{M \setminus i,j}$ and $S_t^{M \setminus i,j}$) are identical under both instruments. Therefore, the difference lies only in the financial payments that are made under both scheme:

$$\Delta_{tax,T}^{M,i} - \Delta_{liability,T}^{M,i} = E_T \left\{ \sum_{t=T}^{+\infty} \beta^{t-T} \left[-\tau_t e_t^{M,i} + \lambda_t S_t^{M,i} \right] \right\} \tag{21}$$

$$= E_T \left\{ \sum_{t=T+1}^{+\infty} \beta^{t-T} \gamma^{t-T} \lambda_t S_T^{M,i} \right\} \tag{22}$$

$$= E_T \left\{ \sum_{t=T+1}^{+\infty} \beta^{t-T} \gamma^{t-T} \lambda_t \right\} S_T^{M,i} \tag{23}$$

$$= \beta \tau_{T+1} \gamma S_T^{M,i} \tag{24}$$

$$= \left(\tau_T - \lambda_T \right) S_T^{M,i} \tag{25}$$

where (22) follows from by Expression (17). The difference is positive, implying that defaulting on climate debt is more tempting (i.e., less costly) than abandoning the Pigovian tax. Notice that the magnitude of the difference, (24), is equal to the present value of taxing today's carbon stock tomorrow. Alternatively, (25) expresses it as the difference between the Pigovian fee associated to emitting today's stock, $S_T^{M,i}$, and the liability fee associated to the same stock.

A.1.2 On the relative temptation to delay the policy

On the other hand, an advantage of the liability scheme is that it is almost costless to enter. This makes the liability scheme more likely to be adopted in the short run than an ex ante policy. To be precise, we compare the net benefits of both schemes over the first L periods—which is equivalent to the cost of postponing for L periods—computed from the point of view of date zero:[32]

$$\begin{aligned}
\Delta^{M,i}_{liability-tax,L} &= E_0 \left\{ \sum_{t=0}^{L-1} \beta^t \left[b^i_t \left(e^{M,i}_t \right) - \lambda_t S^{M,i}_t + \lambda_t \left(\sum_{j \in M} S^{M,j}_t \right) \frac{h^i_t \left(S^M_t \right)}{\sum_{j \in M} h^j_t \left(S^M_t \right)} - h^i_t \left(S^M_t \right) \right] \right. \\
&\left. - \sum_{t=0}^{L-1} \beta^t \left[b^i_t \left(e^{M,i}_t \right) - \tau_t e^{M,i}_t + \lambda_t \left(\sum_{j \in M} S^{M,j}_t \right) \frac{h^i_t \left(S^M_t \right)}{\sum_{j \in M} h^j_t \left(S^M_t \right)} - h^i_t \left(S^M_t \right) \right] \right\} \qquad (26) \\
&= E_0 \left\{ \sum_{t=0}^{L-1} \beta^t \left[-\lambda_t S^{M,i}_t + \tau_t e^{M,i}_t \right] \right\}, \qquad (27)
\end{aligned}$$

where the second equality comes from the fact that both policies implement the same emissions path.

Using again Expression (17), we get:

$$\begin{aligned}
\Delta^{M,i}_{liability-tax,L} &= \beta^{L-1} E_0 \left\{ \beta^{-L+1} \sum_{t=0}^{L-1} \beta^t \left[-\lambda_t S^{M,i}_t + \tau_t e^{M,i}_t \right] \right\} \qquad (28) \\
&= \beta^{L-1} E_0 \left\{ \sum_{t=L}^{+\infty} \beta^{t-L+1} \lambda_t \gamma^{t-L+1} S^{M,i}_{L-1} \right\} \qquad (29) \\
&= \beta^{L-1} E_0 \left\{ \beta\gamma \sum_{t=L}^{+\infty} \beta^{t-L} \gamma^{t-L} \lambda_t \right\} S^{M,i}_{L-1} \qquad (30) \\
&= \beta^L \tau_L \gamma S^{M,i}_{L-1} \qquad (31) \\
&= \beta^{L-1} \left(\tau_{L-1} - \lambda_{L-1} \right) S^{M,i}_{L-1} \qquad (32)
\end{aligned}$$

The sign of the above difference is positive, implying that the liability scheme is strictly less costly over any finite horizon. Its magnitude is the (discounted) expected cost of the inherited stock at date L. It can also be interpreted as the present value, from the point of view of date zero, of the difference between the Pigovian tax associated with emitting the stock at period $L-1$ and the liability associated with that same stock.

Despite the compounded discount factor, β^L, the above difference is not necessarily negligible, even if L is large. In fact, when damage is a convex function of total stock, and stock increases over time, the tax rate $\tau_L = E_L \left[\sum_{t=L}^{+\infty} (\gamma\beta)^{t-L} \frac{dH_t}{dS_t} \right]$ increases with L.[33] Therefore, the size of the difference can even increase with

[32]To focus on the temporal differences in the two schemes, we shall again assume that all tax revenues from are devoted to compensating participating countries in proportion to harm, exactly as under the liability scheme.

[33]As mentioned in the text, the fact that the social cost of carbon increases over time is

L if $\tau_{L+1}/\tau_L > 1/\beta$. With discount factors close to one, this is a distinct possibility.

A.1.3 Comparison

Evaluated from the point of view of date zero, we obtain from Expression (24) that the expected relative temptation to default on liability payments at date T writes:

$$\beta^T E_0\left[\Delta_{tax,T} - \Delta_{liability,T}\right] = \beta^{T+1}\tau_{T+1}\gamma S_T^{M,i} > 0 \tag{33}$$

Comparing with Expression (31), we see that the relative temptation to default on climate debt mirrors the relative temptation to delay the implementation of a Pigovian tax.

A.2 Proof of Proposition 4

Consider a given liability scheme $\{\lambda_t\}_t$. Ignoring the liability market for the time being, the benefits of continued participation at date T to the liability scheme write as follows:

$$\Pi_T^{M,i} = E_T^i\left\{\sum_{t=T}^{+\infty}\beta_i^{t-T}\left[b_t^i\left(e_t^{M,i}\right) - \lambda_t S_t^{M,i} + \lambda_t\left(\sum_{j\in M} S_t^{M,j}\right)\frac{h_t^i\left(S_t^M\right)}{\sum_{j\in M} h_t^j\left(S_t^M\right)} - h_t^i\left(S_t^M\right)\right]\right. \tag{34}$$

$$\left. - \sum_{t=T}^{+\infty}\beta_i^{t-T}\left[b_t^i\left(e_{t,T}^{\emptyset,i}\right) - h_t^i\left(S_{t,T}^{\emptyset}\right)\right]\right\}. \tag{35}$$

where we denote by $e_t^{M,i}$ and $S_t^{M,i}$ the carbon flow and stock of Country i when countries in M implement the liability scheme. Also, $S_t^M = \sum_{j\in N} S_t^{M,j}$. Finally, $e_{t,T}^{\emptyset,i}$ and $S_{t,T}^{\emptyset}$ refer to the carbon flow and stock, respectively, after all countries simultaneously abandon the policy at date T (we are back to the business-as-usual scenario from T onward).

In order to the disentangle the effects of redistribution from the pure benefits of mitigation, we decompose the benefits of continued participation to the liability scheme into two terms:

supported by most integrated assessment models.

$$\Pi_T^{M,i} = \Phi_T^{M,i} + \Gamma_T^{M,i}. \tag{36}$$

The first term is the benefit associated with the emissions reduction that the liability scheme incentivizes:

$$\Phi_T^{M,i} = E_T^i \left\{ \sum_{t=T}^{+\infty} \beta_i^{t-T} \left[b_t^i \left(e_t^{M,i} \right) - h_t^i \left(S_t^M \right) \right] - \sum_{t=T}^{+\infty} \beta_i^{t-T} \left[b_t^i \left(e_{t,T}^{\emptyset,i} \right) - h_t^i \left(S_{t,T}^{\emptyset} \right) \right] \right\}. \tag{37}$$

The second term corresponds to the expected transfers that are made through the compensation of climate harm:

$$\Gamma_T^{M,i} = -E_T^i \left\{ \sum_{t=T}^{+\infty} \beta_i^{t-T} \lambda_t \left[S_t^{M,i} - \left(\sum_{j \in M} S_t^{M,j} \right) \frac{h_t^i \left(S_t^M \right)}{\sum_{j \in M} h_t^j \left(S_t^M \right)} \right] \right\}. \tag{38}$$

Observe that because compensation transfers are balanced at any date we have, for all T and all M:

$$\sum_{i \in M} \Gamma_T^{M,i} = 0. \tag{39}$$

A.2.1 Homogeneous countries

When all countries are homogeneous (same discount factor, same beliefs, same expected harm), no transfers are expected: $\Gamma_T^{M,i} = 0$ for all $i \in M$. Therefore, the existence of the scheme makes Country i better off if and only if $\Phi_T^{M,i}$ is positive.

Consider first the grand coalition ($M = N$). By symmetry, it is plain that $\Phi_T^{N,i} = (1/N) \sum_{j \in N} \Phi_T^{N,j}$. Consider now the liability scheme where carbon stock is priced at current marginal harm, $\lambda_t \equiv H_t'$. Because this liability scheme generates first-best incentives (Expression 7), it maximizes

$$\sum_{j \in N} E_T^i \left(\sum_{t=T}^{+\infty} \beta^{t-T} \left[b_t^j \left(e_t^{N,j} \right) - h_t^j \left(S_t^N \right) \right] \right). \tag{40}$$

Therefore, the sum $\sum_{j \in N} \Phi_T^{N,j}$, which is the net gain of the first-best allocation over the business-as-usual outcome, is strictly positive. It follows that all countries have an interest in participating in the efficient liability scheme.

More generally, for a coalition of M members, there is a net positive gain over the business-as-usual scenario that can be obtained by maximizing

$$\sum_{j\in M}\left(\sum_{t=T}^{+\infty}\beta^{t-T}\left[b_t^j\left(e_t^{M,j}\right)-h_t^j\left(S_t^M\right)\right]\right)$$

over the emissions $\left\{e_t^{M,j}\right\}_{t\geq T}$, for all $j\in M$. This optimum is defined by

$$\frac{db_t^j}{de_t^j}=\sum_{u=t}^{+\infty}(\beta\gamma)^{u-t}\left(\sum_{k\in M}\frac{dh_u^k}{dS_u}\right),$$

for all $t\geq T$ and all $j\in M$ so that it can be implemented by the means of a liability scheme with

$$\lambda_t^M=\sum_{k\in M}\frac{dh_t^k}{dS_t}. \tag{41}$$

To summarize, when countries are symmetric, any coalition M can be sustained for suitable values of the liability rate. In particular, any $\lambda_t\leq\lambda_t^M$ generates welfare gains to every members of the coalition. (By continuity, some values of $\lambda_t>\lambda_t^M$ also generate welfare gains.)

A.2.2 Heterogeneous countries

The mechanism is designed to handle country heterogeneity. It aggregates beliefs and attitudes towards the future (i.e., discounting). Whether it will generate welfare gains for all will depend on how heterogeneous countries are.

Following the above analysis, for any coalition M, there exist values $\lambda_t\leq\lambda_t^M$ of the liability of rates such that there are net collective gains for the coalition as a whole. This implies that *ad hoc* monetary transfers can lead to positive welfare gains for all. However, the mechanism we envision operates transfers based on realized harm, which may not coincide with those transfers necessary to garner participation. When countries are homogeneous this is not a problem, as we have seen, because no transfers are required. Therefore the severity of the participation problem is directly related to the heterogeneity of countries. In fact, when heterogeneity is not too pronounced, no transfer would actually be needed either.

It also follows from the above analysis (Expression (41)) that higher participation allows the liability rate λ_t to become closer to its efficient value λ_t^N, which can always be sustained when all countries participate.

Notice that the larger the participating coalition, M, the larger the benefits from emissions reductions, $\Phi_T^{M,i}$, and the smaller the harm; hence, the smaller the magnitude of the redistributive transfer, $\Gamma_T^{M,i}$. As a result, the more participants, the higher the incentives to participate. This observation hints at the possibility of a positive unraveling effect spurred by an initial "coalition of the willing". Of course, without further assumptions on the heterogeneity of countries, one cannot say for sure how many countries will choose to participate. Clearly, if countries are so heterogeneous that some countries expect no climate harm—or even benefit from climate change—those countries will refuse to participate in a scheme that asks for their contribution but will never benefit them, either through compensation (they will receive none) or by a reduction of the overall externality (they are unaffected by it, or may even benefit from it).

Finally, the above analysis has ignored the fact that liabilities are tradable. Because trade is voluntary, it can only enhance participation.

B Appendix not intended for publication

This appendix shows how incentives are affected when the assumption of a large participating coalition does not hold.

Denote $\mu_i = E_t^i \left[\frac{h_u^i(S_u)}{\sum_{j \in M} h_u^j(S_u)} \right]$, which we will assume to be independent of u (geographical impacts of climate change do not change over time, in expected terms). Then let $\lambda_u = H_u'$. As long as $1-\mu_i$ is not negligible, then $(1-\mu_i)\lambda_u = (1-\mu_i)\sum_{j \in N} \frac{\partial h_u^j}{\partial S_u} >> \frac{\partial h_u^i}{\partial S_u}$. Then Expression (5) becomes:

$$\frac{db_t^i}{de_t^i} = E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \gamma^{u-t} (1-\mu_i) \frac{dH_u}{dS_u} \right\} = (1-\mu_i)\tau_t^i. \tag{42}$$

Trade on the liability market equalizes marginal benefits, as in Expressions (10) and (11):

$$\begin{aligned} e_t^i: \quad \tfrac{db_t^i}{de_t^i} = \quad & E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \gamma^{u-t} \left((1-\mu_i)\lambda_u + \frac{dC_u^i}{dL_u^i} \right) \right\}, && (43) \\ x_t^i: \quad p_t = \quad & E_t^i \left\{ \sum_{u=t}^{+\infty} \beta_i^{u-t} \gamma^{u-t} \left((1-\mu_i)\lambda_u + \frac{dC_u^i}{dL_u^i} \right) \right\}. && (44) \end{aligned}$$

Incentives to reduce emissions are lessened, but this can be addressed by setting a larger liability rate: $\lambda_u > H_u'$.