A liability approach to climate policy: A thought experiment

Etienne Billette de Villemeur and Justin Leroux

December 2016

A liability approach to climate policy: A thought experiment*

Etienne Billette de Villemeur† Justin Leroux‡

*The authors wish to thank Geir Asheim, Hassan Benchekroun, Yann Bramoullé, Alessio D'Amato, Julien Daubanes, Georges Dionne, Paul Ekins, Gérard Gaudet, André Grimaud, Christian Gollier, Jérémy Laurent-Luchetti, Peter Hartley, Marcus Hermann, Louis Hotte, Shi-Ling Hsu, Yann Kervinio, Sonja Köke, Stéphane Lambrecht, Derek Lemoine, Leslie Martin, Stephen Salant, Marc Santugini, Charles Séguin, 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*.

Abstract

We observe that a Pigovian climate policy need not exact full payment of the social cost of carbon upon emission to yield optimal incentives. Following this insight, we propose the creation of a carbon liabilities market to address climate change. Each period, countries would be made liable for their share of responsibility in current climate damage. This yields first-best emissions patterns. Also, because liabilities could be traded like financial debt, it decentralizes the choice a discount rate as well as beliefs about the severity of the climate problem. From an informational standpoint, implementation relies only on realized harm and on the well-documented emission history of countries, unlike a carbon tax or tradable permits scheme, which are based on a sum of discounted expected future marginal damage. We offer a discussion of the differences between a liability scheme and a carbon tax along the dimensions of information, participation, commitment, intergenerational fairness, and exposure to risk.

Keywords: Carbon Liability, Climate Policy, Market Instruments, Pigovian Tax.

JEL classification code: Q54, H23.

1 Introduction

The road to a worldwide agreement to curb climate change is still being paved, most visibly through the emblematic yearly Conferences of Parties (COP) of the United Nations. While the 2015 Paris agreement (COP21)—and its confirmation by COP22—binds participating countries only to a limited extent, it remarkably introduces the notion of "Loss and Damage". In fact, and even though the word is never spelled out, analysts attribute the progress achieved by the COP21 to the shift towards liability considerations (Davenport et al., 2015).

In what follows, we propose a thought experiment to explore how an approach explicitly based on liability could transform climate policy and its corresponding economic instruments. This will require adopting an *ex post* standpoint rather than the *ex ante* approach of all economic instruments introduced thus far. This work is devoted to exploring the pros and the cons of this reversed perspective. We will argue that the latter not only improves the chances of reaching an agreement while promoting first-best emission abatement efforts, but also fosters better intergenerational fairness. To implement this approach, we envision entrusting to an atemporal device (a market for climate debt) the transfer of wealth across generations. A critical issue, however, is the robustness of this very same institution, namely its ability to hold governments accountable, as we shall see.

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 long-term commitment (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 (the *ex post* approach). Economic efficiency, however, is unaffected by the choice of viewpoint. In fact, our thought experiment leads us to the realization that a Pigovian climate policy need not exact the full payment of the social cost of carbon at the time of emission in order to yield first-best incentives, as we now outline.

2 The basic argument

Consider a simple two-period illustration with a single decision maker. She obtains benefit $b(e)$ from emitting e units of emissions in Period 1. A fraction

γ of these emissions will remain in the atmosphere, leading to $h \times (\gamma e)$ units of harm in Period 2. Hence, h can be thought of as the social cost of carbon per unit of greenhouse gas (GHG) residing in the atmosphere. Under an optimal carbon tax, the decision maker is required to pay $\tau = \beta E[\gamma h]$ per unit of e at the time of emission, where β is the discount factor adopted by the tax setter and E is the expectations operator. Consequently, the decision maker chooses her emissions level upon maximizing her total (period-one) payoff:

$$\max_e b(e) - \tau e = \max_e b(e) - (\beta E[\gamma h]) e, \tag{1}$$

so that $b'(e^*) = \beta E[\gamma h]$. If β is actually chosen so as to match the discount factor of society, the resulting emissions level is socially optimal.

Suppose instead that the decision maker is only required to pay the (undiscounted) harm, $\mu \times e = h \times \gamma e$, when it occurs in Period 2. Her objective is now to maximize the discounted sum of her payoffs:

$$\max_e b(e) - \beta E[\mu] e = \max_e b(e) - \beta (E[h\gamma] e),$$

where β now follows from the decision maker's own preferences. Because the objective appears to be unchanged, her emission decision, e^*, is the same as under the optimal carbon tax, because it is such that $b'(e^*) = \beta E[h\gamma]$. If β is equal to the discount factor of society, the resulting emissions level is socially optimal.[1]

Following this insight, we describe what could be a policy instrument: a *carbon liability market.* Upon emitting CO_2 in the atmosphere,[2] countries are issued carbon liability commensurate to their emissions. Liability bearers are legally bound to pay damages over time as climate harm occurs.[3] This effectively requires them to pay for harm 'upon delivery' rather than upfront, as would be the case with a carbon tax. Carbon liability does not expire, but instead decays at the same rate as atmospheric CO_2, all the while holding bearers accountable

[1] Notice, however, that this second approach allows for harm to be observed before payments are made, so that the decision maker never has to pay out-of-pocket amounts linked to harm that is still hypothetical. From the standpoint of political acceptability this may be a key difference, as we shall see.

[2] For expositional purposes, we shall speak only in terms of CO_2.

[3] Clearly, the issue of determining the magnitude of anthropogenic climate harm is a difficult one. Note, however, that this question is by no means resolved with a carbon tax. In fact it is made even worse, because predicting the harm caused by a ton of carbon emitted today decades or centuries into the future is a much more heroic feat than assessing the responsibility of past emissions in (observed) harm occurring today.

for paying carbon damages over time.

Simply put, our proposal consists in converting CO_2 emissions into national (financial) debt. We shall suppose that debt payments are devoted to compensating participating countries to the tune of climate harm incurred.[4] By spreading payments over the long run, a liability scheme avoids the problem that arises when current generations are asked to pay for the sole benefit of future generations, as is the case for all carbon pricing policies proposed to date. To be clear, because individuals and firms are too short-lived to be held accountable in the long run, we suppose that the only entities authorized to hold liability are countries.[5]

Also, carbon liability can be traded on a market. When purchasing extra liability, a country commits to repaying the corresponding climate debt, as future harm occurs, for as long as it bears it. In turn, a seller of liability will be exempt from the associated debt payments. Trade will occur at a price between what the seller is willing to pay and what the buyer is willing to accept to endorse the liability.

To build upon our previous example, consider now that carbon liability is traded on a world market at price p and denote by X the net quantity of liability sold by the decision-making country. Thus, it maximizes $b(e) - \beta E[h\gamma](e - X) - pX - C(p(e - X))$ over e and X, where C is a convex cost function associated with holding climate debt.[6,7] We assume that countries choose to accumulate debt only to the extent that it brings positive net benefit; hence, that $C' < 1$. This yields the following first-order conditions:

$$e: \quad b'(e^*) = \quad \beta E[h\gamma] + pC'(p(e^* - X^*)), \tag{2}$$

$$X: \quad p = \quad \beta E[h\gamma] + pC'(p(e^* - X^*)). \tag{3}$$

Equations (2) and (3) together express the fact that, when faced with liability price p, a country sets its emissions so that its marginal benefit equals this price: $b'(e^*) = p$. From that emission level, whether a country chooses to purchase or

[4]How revenues are used is a question in and of itself. We address this issue in a separate paper (Billette de Villemeur and Leroux, 2011).

[5]We use the word 'country' somewhat loosely. Presumably, nothing prevents the scheme to be applied at a lower level—say at the province or state level in a federal state—provided the corresponding entity is enduring.

[6]Note the minus sign in front of pX. A seller of liability actually *pays* the buyer to hold it in her stead.

[7]In practice, national economies often incur extra costs due to increased borrowing constraints as debt grows (Wachtel and Young, 1987; Engen and Hubbard, 2004; Laubach, 2009).

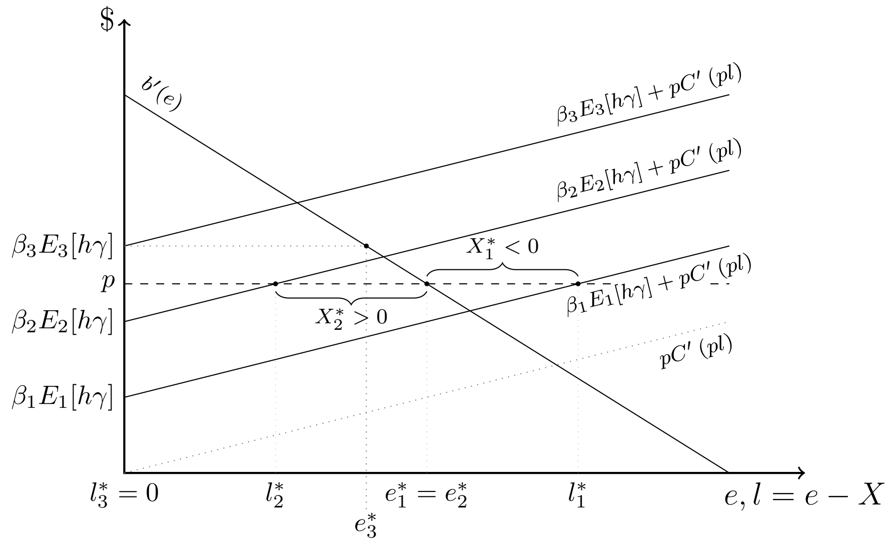

Figure 1: **An individual country's decision.** Facing liability price p, emissions are determined by $b'(e^*) = p$. If the country's discounted value of expectated (global) marginal harm is relatively low ($\beta_1 E_1[h\gamma]$) it will choose to purchase some liability: $X_1^* < 0$. Conversely, if it is relatively high, the country will choose to sell some: $X_2^* > 0$. If it is so high as to lie above the liability price, the country will emit $e_3^* < e_1^* = e_2^*$ such that $b'(e_3^*) = \beta_3 E_3[h\gamma] > p$ and sell all of its liability: $l_3^* = e_3^* - X_3^* = 0$.

sell liability depends on its discounted value of expected global marginal harm (Figure 1). Countries expecting relatively low discounted marginal harm—i.e., countries with a low value of $\beta E\,[h\gamma]$, either because they are skeptical about climate change or because they have little concern for the future—will purchase liability ($X < 0$). They will do so up to the point where the marginal cost of accumulating climate debt compensates the gap between their expectation of discounted marginal harm and the liability price. Likewise, countries with high expected discounted marginal harm will sell some liability.

In equilibrium, the market-clearing price, p, emerges endogenously as the 'global' social cost of carbon (Figure 2). Carbon liability acts as a tradable Arrow-Debreu-type security that makes markets complete, thus yielding allocative efficiency through decentralization. Corner solutions are possible, corresponding to situations where some countries have such a relatively high marginal

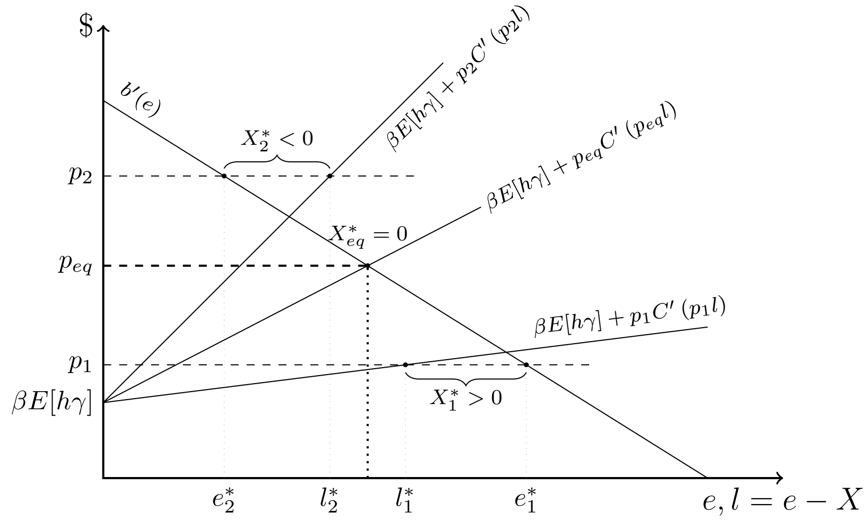

Figure 2: **Liability price determination.** Emissions are decreasing in the liability price: If $p_2 > p_1$, then $e_2^* < e_1^*$ from the downward-sloping marginal benefit curve, $b'(e)$, and the fact that $\beta E[h\gamma] + p_2 C'(p_2 l)$ (which determines l_2^*) is steeper than $\beta E[h\gamma] + p_1 C'(p_1 l)$ (which determines l_1^*). When the price is sufficiently high, there is excess demand of liability on the market ($X_2^* < 0$); conversely, a low price leads to excess supply ($X_1^* > 0$). In equilibrium, the market clears ($X_{eq}^* = 0$).

expected damage—$\beta E[h\gamma]$—that they would sell all of their liability. However, in that case, these countries would actually emit even less than the market price of the liability would imply: for them, $b'(e) > p$ (Country 3 in Figure 1).

Notice that the equilibrium liability price is greater than the expected harm: $p_{eq} \geq \beta E[h\gamma]$ (by Expression 3). This is because of the cost of holding debt, 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, reduces global emissions.

We have laid down the basic logic of a carbon liability scheme. At first sight it appears to be equivalent to a carbon tax. However, climate change is a peculiar externality problem for two reasons. First, it covers an unusually long timespan: some GHGs stay in the atmosphere for hundreds of years (Archer, 2005), with climate consequences to match. Second, the climate problem is surrounded by radical uncertainty, owing to the fact that we may be entering climate regimes yet unseen (IPCC, 2013). This will induce profound differences between the *ex ante* and the *ex post* approach, which deserve further scrutiny. After relating our work to the literature, the bulk of this work will be devoted to discussing the advantages and drawbacks of a liability scheme relative to a carbon tax.

3 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). 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 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).[8] 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 (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 (read liability) is a most natural reflex that should be left unhindered whenever it is an efficient option.

The use of liability 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 mimicking the

[8]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.

physical features of the externality to be managed (Moulin, 2002). 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 the carbon tax and cap-and-trade programs. By making explicit—and, most importantly, financial—the somewhat intangible carbon debt that mankind accumulates along with atmospheric CO_2, carbon liability does just that.

In a climate change context, liability has very recently been proposed as a means to make 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. Ours is mainly one of economic efficiency in that, as compared to a carbon tax, a liability scheme provides the same (first-best) incentives to emit yet makes participation more appealing. This reduces the distortions that usually arise due to limited participation : liability mitigates the externality on a larger scale (more participants) and garners 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 Economic and Social Council, 2016), have flourished over the last 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. One may also wonder about the incentives to keep on devoting resources to generating accurate values for these indicators over the long run if they do not directly affect policy.

By contrast, our proposal consists in providing an indicator of indebtedness that directly impacts policy, because liability implies 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 is the proposal of Gersbach and Winkler (2012) that reverts the Pigovian logic, effectively transforming the carbon tax into an abatement subsidy. Countries would initially contribute large lump sums to an international climate fund and would be refunded over time proportionally to their abatement efforts relative to a business-as-usual scenario. While we retain their use of an international climate fund to collect and redistribute monies to countries, we view the large entry cost as a significant obstacle to their 'global refunding scheme'. An attractive feature of our proposal is that, although generating proper incentives, it reduces substantially the participation cost of countries.

Naturally, estimating harm is central to the climate problem. Therefore, our work is related to the vast literature on the social cost of carbon. The most popular approach in the past decades has been to construct integrated assessment models (IAMs) that combine the dynamics uncovered by both climate science and economics (see Tol, 2009, for a review). While increasingly sophisticated, IAMs are also becoming dramatically less transparent for policy makers. Very recently, research efforts have been made to provide more intelligible estimates of the social cost of carbon. Using a dynamic stochastic general equilibrium model, Golosov et al. (2014) find a formula for the optimal carbon tax that does not rely on future values of output, consumption or stock of CO_2 in the atmosphere. While a remarkable feat, their formula cannot do away with the discount factor, however. Bijgaart et al. (2016) propose a closed-form solution for the social cost of carbon. There, too, knowledge of the discount factor is required. By contrast, an advantage of our approach is that it does not require taking position on the discount factor to compute liability payments, which are only based on realized harm. Yet, for countries to make efficient decisions, they each will have to adopt a discount factor and make predictions about future harm, for which those recent advances will be highly valuable.

4 The pros and cons of a liability scheme

We now discuss the advantages and drawbacks of a liability scheme relative to a carbon tax. Our arguments are organized along the dimensions of information, participation, commitment, intergenerational fairness, and exposure to risk.

4.1 Information

We first compare the carbon liability approach to the carbon tax on informational grounds.

4.1.1 Advantages

Recall that the optimal carbon tax rate is equal $\tau = \beta E[\gamma h]$. We examine each component in turn, keeping in mind that carbon liability payments are determined by $\mu = h\gamma$.

$E[h]$**:** ***Ex post*** **vs.** ***ex ante.*** The carbon tax follows a resolutely *ex ante* approach. In particular, its amount is established on the basis of the expected discounted future harm. By contrast, carbon liability requires payments to be based on realized harm. By nature, it thus inherits all the informational advantages attached to an *ex post* approach. Primarily, the fact that harm is actually realized at the time of payment makes it possible to measure: payments do not depend on hypotheses about the distant future. It also prevents countries from strategically downplaying their expectations about future harm.

Of course, even realized harm is difficult to measure. First, like in any field of empirical research, one may expect studies to lead to a range of estimates rather than to a unique number. Moreover, 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 without anthropogenic emissions. In other words, assessing man-made harm relies on comparing an observable situation to a hypothetical one. While such an endeavor may raise skepticism, it is actually better grounded than estimating the optimal level of a carbon tax, which actually relies on the comparison of scenarios that are *all* hypothetical—and about the distant future, at that.

β **: Objective vs. subjective.** Because the optimal level of a carbon tax reflects expected discounted marginal harm, its computation relies on a subjective parameter, namely the rate of time preference. Yet, eliciting preferences has always been considered a difficult exercise. On may suspect countries to misreport their rate of time preference in order to influence the tax rate in their favor.

By contrast, the implementation of carbon liability 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.

Moreover, the fact that the discount factor does not enter the computation of liability payments actually allows the scheme to accommodate heterogeneous time preferences across countries. As we shall discuss in Section 4.2.1, this eliminates the need for a consensus on the discount factor to be used.

Of course, this greater degree of decentralization (as compared to the carbon tax) 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?

If decision makers display a systematic bias for the present—say, to inflate their performance while in office, for electoral concerns—the carbon liability scheme would fail to implement the optimal emissions path. However, those very same considerations currently lead decision makers to adopt lower-than-optimal carbon tax rates. Actually, this problem does not follow from a difficulty in eliciting preferences. Rather, it is a political economy problem that arises from the fact that decision makers have goals that differ from aiming at reflecting exactly the preferences of their constituency. As we shall see in Section 4.4.2, a carbon liability market better aligns the choices of decision makers with the interests of the population, by making more visible future consequences of current policies.

The real question surrounding a country's choice of β is 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. Along this line, as already mentioned, the carbon liability scheme outperforms the carbon tax in that there is no need to elicit the time preferences of the population: by design, the incentives to the current generation reflect the *exact* value of the discount factor.

$E[\gamma]$**: Tax on stocks vs. tax on flows.** Climate harm, which is stochastic by its very nature, depends on the current concentration—i.e., the stock—of GHGs in the atmosphere. It is thus related to past emission flows, although only indirectly. The relationship between past emission flows and GHG concentration is often expressed by means of a decay rate, which summarizes complex geophysical and atmospheric phenomena and depends notably on GHG concentration (Prather, 1996; IPCC, 2001). Although this relationship is most likely deterministic, an accurate computation of the carbon tax rate still requires i) an exact 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 decay rate can be considered to be unknown to the decision maker, who must rely on an expected value only.

By contrast, carbon liability mimics the physical nature of the externality by directly relating realized harm to the current stock of emissions attributable to each country. It follows that there is no need to construct emissions scenarios because payments rely only on past emissions. The stock attributable to each country is computed from the knowledge of its past emissions and of the decay rate. The former can be obtained from national accounting on the most GHG-intensive industries (fossil fuels, cement, steel, etc.). The latter can be directly computed as the ratio of current GHG concentration (net of current emissions) and previous GHG concentration. All calculations are thus fact based rather than scenario based.

4.1.2 Limits

A displaced informational burden. The informational advantage identified above should not hide the fact that carbon liability displaces the informational burden onto countries. Indeed, whereas the planner need not make predictions about future harm scenarios or pinpoint a discount factor, it is up to the countries themselves to do so. In fact, 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.

Yet, forecasting man-made climate harm need not be a solitary exercise, to be made redundantly by each country. Instead, while countries remain free to reach their own conclusions, their forecasts can result from a cooperative process, or they may gather information from an outside independent agency (the IPCC comes to mind). Furthermore, while each country must adopt a discount factor, it does not need to agree with others on its value under a carbon liability approach. We view this very fact as an advantage rather than a drawback (see Section 4.2.1).

Lessened incentives for adaptation. Because harm is compensated under a carbon liability approach, the latter 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. As a result, insured parties must still face residual risk even though it would have been possible to fully insure them absent this moral hazard problem. Similarly, in order to sustain adaptation efforts, it seems desirable to limit the extent to which climate harm is compensated, and to compensate harm only above a certain threshold (the equivalent of a deductible). This has the advantage of not requiring the evaluation of small harm but only calls for evaluation after occurrences of 'severe' harm, the magnitude of which should be defined collectively.

However, 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.

On top of these considerations, it is worth noting that holding a carbon liability 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 the efficient allocation of adaptation efforts.

4.2 Participation

Many of the advantages of using a carbon liability market over a carbon tax pertain to participation. We discuss them in turn.

4.2.1 Increased decentralization

A clear advantage of the liability approach 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 a carbon liability market. Each country can have its own beliefs about future climate harm and its own attitude towards future generations by choosing the amount of liability it wishes to hold.

This has 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.[9] By reducing the number of dimensions on which countries need to agree, one can expect a carbon liability market to greatly facilitate the negotiation process toward a powerful climate treaty.

4.2.2 Incentives to participate in the market

In addition to enhancing participation by increasing decentralization, a carbon liability market promotes participation from countries that may initially have little interest in a scheme aimed at reducing GHG emissions. This follows from the profit opportunities that comes with participating in any market. In fact, to take an extreme example, even a country that denies the existence of a climate problem will want to participate in a carbon 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 carbon liability market can be expected to garner the participation of many countries.

With a carbon tax, on the other hand, participation is very difficult to obtain from any but the most concerned countries. In particular, for the climate-skeptic country of our example, it is simply a nonstarter.

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

4.2.3 State sovereignty

A difficulty attached to global endeavors is that governments do not like to be bound in their policy choices. As such, a global carbon tax can be considered as very intrusive. By contrast, a carbon liability approach offers countries more freedom in how to tackle the climate issue. In fact, a carbon liability converts 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, a tradable allowance scheme, or any other instrument (or mixture of instruments) of that country's choice. State sovereignty is thus a further argument towards the adoption of a liability approach.

4.2.4 Payments are spread over time and match realized harm

A difficulty attached to the carbon tax is that it demands emitters to pay immediately for all subsequent expected damage. This creates significant hurdles for participation for two reasons. First of all, it requires emitters to pay 'up front' for the entire future harm attached to their emissions. But also, it makes emitters pay for harm that is still uncertain, and which may or may not materialize in the end.

Carbon liability does not suffer from either of these drawbacks because they require emitters to pay over time as harm occurs. By delaying payment, the initial phase of a climate agreement is a smooth process rather than a shock therapy. More importantly, carbon liability payments match realized harm. This has two important advantages. For one thing, carbon liability does not involve out-of-pocket payments if no harm occurs. For another, if realized harm exceeds expected harm, carbon liability guarantees that sufficient funds will be collected to compensate those that end up suffering from climate harm. In other words, being an *ex post* instrument, carbon liability can better handle distributional concerns, both with respect to emitters—who will not be unduly solicited—and with respect to victims of climate harm—who will not be left uncompensated. These are clear advantages over the carbon tax that make carbon liability more likely to rally supporters.

4.2.5 Incentives to adapt, which lower the overall cost of climate change

The carbon tax, by its *ex ante* nature, disregards entirely the actual realization of harm. As a result, there are no incentives for emitters to make adaptation efforts outside of its borders because each country is left alone in the face of the consequences of climate change. By making adaptation beneficial to all countries holding liability, a carbon liability market turns adaptation into a public good. Like all public goods, it is subject to the well-known problem of under-provision. However, it is in the interest of all countries to put in place institutions or mechanisms to overcome this problem.

One possibility would be to keep track of the adaptation efforts of each country and credit them for a substantial part of the returns (i.e. avoided harm). This privatization of sorts of the returns to adaptation would provide efficient incentives for countries to invest in adaptation and to seek out the most profitable opportunities around the globe. It would also provide a means for countries to hedge their liability positions against climate risk.

The incentives to make adaptation efforts that are attached to a carbon liability market will translate into lower expected harm. As a result, even in expected terms, an *ex post* approach is likely to be cheaper than an *ex ante* one, and hence easier to adopt.

4.3 Commitment and credibility

This section examines the liability approach through the lens of commitment, which is certainly the most controversial aspect of our proposal.

4.3.1 Unconcern for future generations and strong incentives for renegotiation (but not worse than postponing)

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

Across generations: Political short-termism First, the relatively short-sighted nature of politicians, due to short term lengths, may induce them to neglect the burden they impose upon their successors. As a result, and although they may be eager to participate in a carbon liability market (as seen in Section 4.2.2), governments may tend to underestimate the (future) cost of emitting (today) and be inclined to allow for more emissions than would be in the country's (intertemporal) interest. One can expect this lack concern for future generations to undermine the efficiency of a carbon liability market.

However, the carbon tax is not immune to the shortsightedness of politicians, either. The latter simply takes on a different form, whereby the implicit price of carbon in effective tax rates—and cap-and-trade prices—are widely considered by experts to be far below most reasonable estimates of the social cost of carbon. Moreover, by making visible the debt left to future generations a liability approach appears better equipped to fight against the shortsightedness of governments. It does so by generating an indicator—financial climate debt—that civil society can use to monitor the decisions of its rulers (see Section 4.4.2).

Across countries: Temptation to default on liability payments The second problem arises between liable countries and those who are to be compensated. Indeed, as time goes on and as the burden increases over time, one may expect liable countries to choose to default on their liability payments. This could leave countries affected by climate change without the compensations they were promised, with possibly dramatic consequences to their populations. This is a legitimate concern, which should not be ignored. That being said, a carbon liability market actually provides an incentive device towards honoring liability payments through the 'financialization' of climate harm. Indeed, defaulters would not simply have to bear the moral consequences of reneging on their commitment, but would also suffer financial setbacks: Under a carbon liability scheme, failure to honor liability payments is no different than defaulting on financial debt.

From a moral point of view, however, there is a difference in the nature of these debts. In the case of default on financial debt, those affected are creditors who had chosen to expose themselves to financial risk. By contrast, defaults on liability payments are likely to impact the residents of vulnerable countries that did not choose to expose themselves to climate risk.[10] This makes the problem of defaulting of particular prominence from a moral standpoint.

This temptation is no worse than that of postponing the carbon tax By definition, the carbon tax raises no such commitment issues because payments are made at the time of emission. Nevertheless, instead of the issue of

[10] We thank Yann Kervinio for this observation.

honoring payments, it is the very willingness to participate in the mechanism that is fragile. The issue there is a desire for politicians of emitting countries to postpone the implementation of a carbon tax in order to enjoy the support of high-emitting industries while offering voters immediate economic rewards. Whether this political-economy problem is more or less serious for the carbon tax than for carbon liability is an open question. We show in Appendix A that the temptation to default on liability payments actually mirrors the temptation to continually delay the implementation of a carbon tax. In our opinion, this tradeoff between abandoning a carbon liability scheme in the future and postponing a carbon tax today leans towards a market for liability being the more effective policy. Indeed, before possibly abandoning the program, at least some emissions reduction is being achieved.

Interaction between incentives and participation/commitment. The political-economy problems just discussed unfold along two different dimensions. The intergenerational problem yields low incentives to reduce emissions, under both a carbon tax and a carbon liability market. The across-country problem yields a participation/commitment dilemma: participation for the carbon tax (temptation to delay indefinitely) and commitment for the carbon 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. On the other hand, delayed participation would make any given carbon tax rate even lower relative to the optimal one.[11] Summing up, with a carbon tax, the incentives problem alleviates the participation problem whereas the latter reinforces the former. Similarly, the incentives and commitment problems are related for a carbon liability market, although somewhat differently. A lack of concern for future generations yields too much CO_2 to be emitted, which exacerbates the commitment problem (due to holding too much 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 drawback of a carbon liability market.

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

4.3.2 Market design

Following the previous discussion, it becomes apparent that a carbon liability market requires strong institutions in order to function properly. Implementing a carbon liability market involves five major steps: tracking the emissions of countries, attributing liability to emitters, assessing harm, imposing liability payments, and managing the transfer of liability within the market. Hence, the mechanism minimally requires a central authority. However, several of the above-mentioned duties are already handled by well-established organizations. Reliable databases exist for historical emissions. Assessing their role in current GHG concentration and climate conditions is the focus of many reputable research centers on climate science. Insurance companies frequently assess damage including when linked to climate events. Most countries routinely honor their debt. Finally, government bonds are traded daily on international financial markets.

Regardless, one may still question the credibility of such a mechanism and in particular its supporting central institution. We argue that the fact that the latter supports a market actually makes it relatively robust. Indeed, while all markets require a minimum of supporting institutions, they are rarely called into question for the purported frailty of these very same institutions. We believe that turning the climate externality into a financial entity is actually a source of robustness. It allows the climate problem to transcend mere ecological concerns, which can easily be underrepresented on the political stage. This makes the climate problem tangible and more difficult to ignore or downplay.

4.4 Intergenerational fairness

Although our analysis is mainly focused on economic efficiency, we shall observe that a carbon liability scheme brings significant benefits in terms of intergenerational fairness.

4.4.1 Historical emissions and the distribution of the (financial) burden

An important question when addressing the issue of intergenerational fairness is how to consider emissions that occurred before awareness of the climate change problem. In law theory, a general principle of justice is that no party can be liable in the absence of 'constructive notice'.[12] In other words, a country should not be asked to pay for the climate consequences of emissions made before the discovery of the impact of greenhouse gases on the climate (the year 1990, which saw the release of the IPCC's first assessment report on climate, is usually considered a landmark year). On the other hand, climate change (including currently observed impacts) is mainly the consequence of past actions, which justifies taking into account the historical process leading to its occurrence.

This question highlights further differences between the carbon tax and a carbon liability scheme. They are normative, related to incentives, and political.

Normative By construction, a tax abides by the legal principle of constructive notice. It does so because it applies only to current emissions, and at a time when no country can reasonably ignore that emissions have long-lasting climate effects. Yet, implicitly, the carbon tax completely absolves past emitters from the consequences of their past deeds, and places the burden entirely on current emitters.

By contrast, a carbon liability scheme requires taking a stance on a date from which countries can be considered responsible for their emissions. This choice 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 this choice. Ultimately, unlike a carbon tax, a carbon liability scheme can handle this important distributional concern. Furthermore, it does so in a transparent way 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.

Incentives/efficiency The key feature of a carbon liability mechanism is that the prospect of being liable for future damage 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.[13] As mentioned above, this allows the mechanism to comply with

[12]We thank Shi-Ling Hsu for bringing this issue to our attention.

[13]Various choices of a starting date effectively correspond to different sets of lump-sum transfers between countries.

the basic legal principle of constructive notice while providing a tool to address the normative role of historical emissions.[14]

By contrast, a carbon tax suffers from the usual tension between efficiency and redistribution. If set at the optimal level, so as to implements the optimal emissions pattern, it is a regressive instrument because it is disproportionately generous towards past emitters, who also happen to be today's richest countries. Any departure from this rate, while possibly having positive distributional impacts, will inevitably undermine the efficiency of the tax.

Political Because of this last aspect, a carbon liability scheme may actually ease international negotiations. Indeed, it allows countries to address issues of fairness without impinging on efficiency. By contrast, climate negotiations associated with a carbon tax—or, more generally, with emission targets—necessarily trade off fairness for efficiency, which leads to compromises that are deemed unacceptable by some on efficiency grounds and by others on distributive ones. The coalition formation generated by the Kyoto Protocol (developed countries on one side and developing countries on the other) is telling.

From a political point of view, an advantage of the carbon 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 carbon 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. Hence, responsibility is no longer binary.

Moreover, fine-tuning the starting date allows for smooth adjustments in the responsibility borne by countries. The earlier this date, the higher the burden to more developed countries, as per the high correlation of past emissions with current GDP levels. Finally, 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

[14] 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 damage.

of belonging or not to a category (i.e., located close to a threshold). Next, the adjustments in the starting date affects countries gradually, as opposed to having decisive consequences to those same marginal countries whereas others are not affected at all. In terms of climate agreements, negotiations about the starting date will re-balance the stakes and spread them from a few countries to all nations instead of entrenching the positions of opposed coalitions.

4.4.2 Increased government accountability

A carbon liability scheme converts into financial debt what current generations owe to future generations for the climate harm that their emissions will induce. By making this debt financial, rather than environmental, it becomes both explicit and more difficult to ignore.

It becomes explicit because the financial value of the debt is indeed a market assessment of future harm. The market (liability) price aggregates all available private information and makes it public. Thus, the total liability value at any current date is the best possible estimate of the cost of future climate harm.

In addition, the debt to future generations also becomes more difficult to ignore. An environmental debt remains a moral debt, on which default has few consequences for the debtor. By contrast, defaulting on financial debt brings about tangible consequences to a country's economy, in particular through reduced borrowing and trade opportunities.

4.4.3 Making provisions for the future

A consequence of the financial nature of the debt and the induced increased government accountability is that governments have no choice but 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 capital. As we saw in Section 4.2.3, 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 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 carbon liability).

One may question the ability of governments to make provisions for the future. For example, several countries already face difficulties in honoring their promises to retirees. 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 no tools 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 is tantamount to a "pay-as-you-emit" scheme, whereas a carbon liability market resembles a fully funded system, where liability returns are studied closely by financial markets.

4.5 Exposure to risk

Lastly, an examination of climate policy would not be complete without discussing the risk dimension, which we do now.

4.5.1 Addressing inequalities in realized climate harm

A central feature of the liability approach is its *ex post* nature: its design is based on realized uncertainty. By contrast, a carbon tax embodies an *ex ante* viewpoint: what matters are the expected consequences of emissions so as to induce emitters to internalize climate harm in their decision making. The realized distribution of harm is of no importance because it does not affect efficiency. However, if realized inequalities have any value, a liability approach is better by construction. In terms of incentives, it induces the same emissions allocation, but in terms of distribution, it strictly dominates the carbon tax by embedding compensations to countries actually affected by climate harm.

4.5.2 Risk is borne by those who are responsible for it

With a carbon tax, the emitting generation is completely absolved of the consequences of its emissions upon paying the tax. The risk of climate damage imposed by the accumulated emissions is borne by the most vulnerable countries, regardless of responsibility. It is one thing to share risk among countries, but quite another to be fully exonerated of the risk imposed onto others. Given the magnitude of the possible harm, full exoneration seems thoroughly unreasonable.

By contrast, with carbon liability, risk is now borne by those who are responsible for its occurrence. Because liability payments are used to compensate realized harm, it is the emitting countries that are held accountable for the climate risk—to the tune of their cumulative emissions—rather than those exposed to climate damage. When liability is traded, the purchasing countries will end up bearing the risk. However, not only will they have chosen to do so, but they will have been compensated for doing so by financial gain through the liability price.

4.5.3 Varying debt payments

A possible objection to the carbon liability approach is that it exacts varying debt payments from liable countries. This follows from the fact that the magnitude of the payments is determined by the magnitude of realized climate harm. In fact, with a carbon liability scheme, countries are not contracting on monetary amounts determined *ex ante* but on a share of future, but also uncertain, climate harm.

By contrast, a carbon tax has the apparent advantage of exacting payments that are known at the time of emission. Thus, it facilitates the decision process of emitters. However, this advantage hides the fact that climate harm is inherently fluctuating and unpredictable, so that realized harm almost always differs from expected harm. The difference between actual harm and expected harm is effectively paid by the emitters themselves—should the realized harm fall short of expectations—or, worse, by those countries who bear the brunt of the consequences in the form of (unpredicted) harm—should the realized harm exceed that conveyed by the tax rate.

Under a liability scheme, those variations would be fully compensated by liability payments. More generally, a liability scheme translates into financial terms the very same risk (although environmental, not financial) that countries face under a carbon tax. In other words, it does not introduce risk; rather, it makes explicit risk that was already there.

5 Conclusion

We have proposed a new approach to climate policy based on the trading of carbon liability between countries and have demonstrated that it yields the same emissions patterns as the optimal carbon tax.

Although it adopts an *ex post* approach, unlike traditional economic instruments, our proposal 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. This increased decentralization explains the significant advantages of carbon liability over the carbon tax: it is less informationally demanding, it allows for disagreements in preferences and beliefs through decentralization, and it exacts payments commensurate to *actual* climate damage.

In addition, a liability scheme solves an important problem of climate justice, which is that upon implementing a carbon tax, or any other *ex ante* instrument, emitting countries are absolved from all climate risk. By contrast, a liability scheme holds countries (financially) accountable for the climate harm borne of their emissions over time. This gives governments an incentive to take seriously the climate change problem and to take proactive measures.

Our proposal raises at least one important political-economy concern, however. Countries may find it tempting to abandon the carbon liability program in the future as payments increase due to climate harm becoming more pronounced over time. Problems of incentives to renegotiate should be considered explicitly in this context, as done in Harstad (2012). We leave this for future research.

Another question that needs to be addressed explicitly is that of country participation. The expected net discounted payoffs of countries are identical under a carbon liability scheme and under a carbon tax. Adopting a dynamic framework, we have seen that participation to a liability scheme is cheaper than participation to a carbon tax. However, as made evident by the literature on international environmental agreements and, in particular, by the literature on coalition formation (Barrett, 1994; Bloch, 1996; Diamantoudi and Sartzetakis, 2015), participation of one country depends on that of others. Therefore, the question of participation should be addressed explicitly in future work, for example along the lines of Martimort and Sand-Zantman (2015).

That being said, the liability approach reduces the number of dimensions of potential disagreement between would-be participants. Indeed, agreeing on a carbon tax requires agreement on the discount factor and on the magnitude of yet unrealized climate events far off in the future, two highly debated issues that are unlikely to garner consensus anytime soon. By contrast, the liability scheme 'only' requires agreement on a starting date from which emissions should be counted. While such an agreement is by no means guaranteed, as there is a clear tension between developed and developing countries in this regard, reducing the dimensions of potential disagreement may significantly improve the likelihood of a powerful joint international effort.

References

[1] Archer, D. (2005) "Fate of Fossil Fuel CO2 in Geologic Time," *Journal of Geophysical Research*, **110**, doi :10.1029/2004JC002625.

[2] Barrett, S. (1994) "Self-Enforcing International Environmental Agreements," *Oxford Economic Papers*, **46**, Special Issue on Environmental Economics, 878-894.

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

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

[5] Bloch, F. (1996) "Sequential formation of coalitions in games with externalities and fixed payoff division," *Games and Economic Behavior*, **14**, 90-123.

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

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

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

[9] Diamantoudi, E. and E.S., Sartzetakis, (2015) "International environmental agreements: coordinated action under foresight," *Economic Theory*, **59**, 527-546.

[10] 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.

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

[12] 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.

[13] 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.

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

[15] Golosov, M., Hassler, J., Krusell, P. and A. Tsyvinski (2014) "Optimal taxes on fossil fuel in general equilibrium," *Econometrica*, **82**(1) 41–88.

[16] Hardin, G. (1968) "The Tragedy of the Commons", *Science*, **162**, 1243-1248.

[17] Harstad, B. (2012) "Climate Contracts: A Game of Emissions, Investments, Negociations and Renegociations," *Review of Economic Studies*, **79**, 1527-1557.

[18] Hartwick, J. M., (1990) "Natural resources, national accounting and economic depreciation," *Journal of Public Economics* **43**(3), 291-304.

[19] 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.

[20] 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.

[21] Hsu, S.-L. (2011) *The Case for a Carbon Tax*, Island Press eds.

[22] 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)

[23] IPCC [2013] : Climate Change 2013 : The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, https://www.ipcc.ch/report/ar5/wg1/ (accessed Dec. 7th 2016)

[24] 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.

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

[26] Martimort, D. and W. Sand-Zantman (2016) "A Mechanism Design Approach to Climate Agreements," *Journal of the European Economic Association*, **14(3)**, 669-718.

[27] 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.

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

[29] US 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)

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

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

[32] Tol, R.S.J. (2009) “The Economic Effects of Climate Change,” *The Journal of Economic Perspectives*, **23(2)**, 29-51.

[33] 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)

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

[35] 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)

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

A Appendix

This appendix formalizes the discussion introduced in Section 4.3.1 about the temptation of countries to default on carbon debt. Section A.1 lays down the formal setup, which extends that of the Introduction to a multi-period framework. Section A.2, shows that the temptation to default actually mirrors that of postponing the implementation of a carbon tax.

A.1 The formal setup

In order to make explicit the statement of Section 4.3.1, we present a richer model than the one in the introduction. Denote by h_t^i the function representing the flow of stochastic harm borne by Country i at period t. This harm depends on the current stock of anthropogenic CO_2 in the atmosphere, Z_t, which is itself the sum of the stock of CO_2 that can be attributed to each country: $Z_t^i = \sum_{s=-\infty}^{t} \gamma^s X_s^i$ (it is the sum of its past emissions X_s^i, for all $s \leq t$, accounting for the decay rate γ).[15,16] The global flow of harm at period t is the sum of the harm borne by individual countries: $h_t(Z_t) = \sum_i h_t^i(Z_t)$.

With this notation, the so-called Pigovian tax rate writes:

$$\tau_t \equiv E_t \left[\sum_{s=t}^{+\infty} \beta^{s-t} \frac{dh_s}{dX_t} \right], \tag{4}$$

the discounted sum of expected future marginal harm.

Our proposal consists in converting CO_2 emissions into financial debt. More precisely, in each period, Country i is required to contribute to an international climate fund to the tune of $\mu_t Z_t^i$ where

$$\mu_t \equiv dh_t/dZ_t \tag{5}$$

is the (current) marginal climate harm due to anthropogenic emissions.[17]

Throughout, we shall work under the usual Pigovian assumption that strategic issues associated with the relative sizes of countries can be ignored. Formally, we assume that:

$$\frac{d\mu_s}{dX_t^i} = \frac{d\tau_s}{dX_t^i} = 0 \qquad \forall t \leq s \quad \forall i, \tag{6}$$

which implies that no country can affect the emissions decision of other countries through its impact on the liability payments (or the tax rate).

Moreover, we assume for expositional purposes that domestic marginal harm is negligible compared to global marginal harm:

$$dh_t^i/dZ_t \ll dh_t/dZ_t \qquad \forall i \forall t, \tag{7}$$

which implies that none of the climate harm is internalized in the absence of a policy instrument.

While not completely realistic, these assumptions makes it easier to compare our proposal to the usual schemes—namely, the Pigovian carbon tax and cap-

[15]We could allow for the decay rate, γ, to vary with concentrations of CO_2, but refrain from doing so for the sake of readability.

[16]As mentioned, 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 (developed) countries while preserving incentives at the margin.

[17]Unlike in tort law, we do not aim for "full liability" because, from an efficiency standpoint, it is not optimal to cover the total harm.

and-trade solutions—, which have been derived under the same assumptions, and allows us to focus on the general features of the mechanism.

Under our assumptions,

$$\frac{dZ_s}{dX_t} = \gamma^{s-t} \qquad \forall t \leq s, \tag{8}$$

so that

$$\frac{dh_s}{dX_t} = \gamma^{s-t} \frac{dh_s}{dZ_s} \qquad \forall t \leq s \tag{9}$$

and, in turn,

$$\tau_t = E_t \left[\sum_{s=t}^{+\infty} (\gamma\beta)^{s-t} \frac{dh_s}{dZ_s} \right]. \tag{10}$$

Under a carbon tax, Country i evaluates its present expected net benefit as:

$$PENB_t^i = \sum_{s=t}^{+\infty} \beta^{s-t} E_t \left[b_s^i \left(X_s^i \right) - h_s^i \left(Z_s \right) - \tau_s X_s^i \right], \tag{11}$$

and chooses an emissions stream $\left\{ X_t^i \right\}_{t=0}^{+\infty}$ such that:

$$\frac{db_t^i}{dX_t^i} = \sum_{s=t}^{+\infty} \beta^{s-t} E_t \left[\frac{dh_s^i}{dZ_s} \frac{dZ_s}{dX_t^i} + \frac{d\tau_s}{dX_t^i} X_s^i + \tau_s \frac{dX_s^i}{dX_t^i} \right] \tag{12}$$

$$= \left\{ \sum_{s=t}^{+\infty} \beta^{s-t} E_t \left[\frac{dh_s^i}{dZ_s} \gamma^{s-t} \right] \right\} + \tau_t \tag{13}$$

$$= \left\{ \sum_{s=t}^{+\infty} (\beta\gamma)^{s-t} E_t \left[\frac{dh_s^i}{dZ_s} + \frac{dh_s}{dZ_s} \right] \right\} \tag{14}$$

$$\approx \left\{ \sum_{s=t}^{+\infty} (\beta\gamma)^{s-t} E_t \left[\frac{dh_s}{dZ_s} \right] \right\} = \tau_t, \tag{15}$$

where the second line follows from Expression (8), Assumption (6), and the fact that emissions across periods are independent decisions, so that $dX_s^i/dX_t^i = 0$ whenever $s > t$. The third line follows from the expression of the tax rate (Expression 10) and the fourth from Assumption (7), which states that no country is large enough to internalize a significant portion of the global harm.

Under a carbon liability scheme, Country i evaluates its present expected net benefit as:

$$PENB_t^i = \sum_{s=t}^{+\infty} \beta^{s-t} E_t \left[b_s^i \left(X_s^i \right) - h_s^i \left(Z_s \right) - \mu_s Z_s^i \right], \tag{16}$$

and then chooses an emissions stream $\left\{ X_t^i \right\}_{t=0}^{+\infty}$ such that:

$$\frac{db_t^i}{dX_t^i} = \sum_{s=t}^{+\infty} \beta^{s-t} E_t \left[\frac{dh_s^i}{dZ_s} \frac{dZ_s}{dX_t^i} + \frac{d\mu_s}{dX_t^i} Z_s^i + \mu_s \frac{dZ_s^i}{dX_t^i} \right] \tag{17}$$

$$= \sum_{s=t}^{+\infty} \beta^{s-t} E_t \left[\frac{dh_s^i}{dZ_s} \gamma^{s-t} + 0 + \mu_s \gamma^{s-t} \right] \tag{18}$$

$$= \sum_{s=t}^{+\infty} \left(\beta\gamma \right)^{s-t} E_t \left[\frac{dh_s^i}{dZ_s} + \frac{dh_s}{dZ_s} \right] \tag{19}$$

$$\approx \left\{ \sum_{s=t}^{+\infty} \left(\beta\gamma \right)^{s-t} E_t \left[\frac{dh_s}{dZ_s} \right] \right\} = \tau_t. \tag{20}$$

Again, the second line follows from Expression (8) and Assumption (6). The third line follows from the expression of the liability payment (Expression 5) and the fourth from Assumption (7). The comparison of (14) and (19) makes it clear that even without Assumption (7), both instruments decentralize the very same allocation.[18]

A.2 Defaulting on carbon debt vs. delaying the carbon tax

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 the mechanism and the relative temptation to delay the implementation of both mechanisms. We shall assume throughout that the decision of a country does not affect the behavior of other countries. This assumption is in line with the Pigovian assumptions (6) and (7) above.

[18] 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.

A.2.1 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 carbon debt. On the other hand, because climate damage is likely to increase over time, there is also an increasing temptation to not participate in a carbon tax.

The expected cost of abandoning the carbon tax at a given date T is equal to:

$$\Delta_{tax,T} = E_T \left\{ \sum_{t=T}^{+\infty} \beta^{t-T} \left[b_t^i \left(X_t^i \right) - \tau_t X_t^i - h_t^i \left(Z_t \right) \right] - \sum_{t=T}^{+\infty} \beta^{t-T} \left[b_t^i \left(\tilde{X}_{t,T}^i \right) - h_t^i \left(\tilde{Z}_{t,T} \right) \right] \right\}, \tag{21}$$

where X_i and Z_i (respectively, $\tilde{X}_i$ and $\tilde{Z}_i$) refer to the carbon flow and stock in the case where country i is taxed (respectively, no longer taxed).
Likewise, the expected cost of defaulting on carbon debt from date T onward is equal to:

$$\Delta_{liability,T} = E_T \left\{ \sum_{t=T}^{+\infty} \beta^{t-T} \left[b_t^i \left(X_t^i \right) - \mu_t Z_t^i - h_t^i \left(Z_t \right) \right] - \sum_{t=T}^{+\infty} \beta^{t-T} \left[b_t^i \left(\tilde{X}_{t,T}^i \right) - h_t^i \left(\tilde{Z}_{t,T} \right) \right] \right\}. \tag{22}$$

The rest of the argument will make use of the fact that X_i and Z_i are identical under both instruments (recall Expressions (15) and (20)). Moreover, this implies that if the country abandons the policy, the subsequent emissions flows and stocks, $\tilde{X}_i$ and $\tilde{Z}_i$), will continue to be identical. Therefore, the difference lies only in the financial payments that countries make under both schemes:

$$\Delta_{tax,T} - \Delta_{liability,T} = -E_T \left[\sum_{t=T}^{+\infty} \beta^{t-T} \tau_t X_t^i \right] + E_T \left[\sum_{t=T}^{+\infty} \beta^{t-T} \mu_t Z_t^i \right] \tag{23}$$

The financial gain from reneging on the carbon tax at date T onward is given by:

$$E_T\left[\sum_{t=T}^{+\infty}\beta^{t-T}\tau_t X_t^i\right] = E_T\left\{\sum_{t=T}^{+\infty}\beta^{t-T}\frac{d}{dX_t^i}\left\{E_t\left[\sum_{s=t}^{+\infty}\beta^{s-t}h_s\left(Z_s\right)\right]\right\}\times X_t^i\right\} \tag{24}$$

$$= E_T\left\{\sum_{t=T}^{+\infty}\beta^{t-T}\frac{d}{dX_t^i}\left[\sum_{s=t}^{+\infty}\beta^{s-t}h_s\left(Z_s\right)\right]\times X_t^i\right\} \tag{25}$$

$$= E_T\left\{\sum_{t=T}^{+\infty}\sum_{s=t}^{+\infty}\beta^{s-T}\frac{d}{dX_t^i}\left[h_s\left(Z_s\right)\right]\times X_t^i\right\} \tag{26}$$

$$= E_T\left\{\sum_{s=T}^{+\infty}\sum_{t=T}^{s}\beta^{s-T}\frac{d}{dX_t^i}\left[h_s\left(Z_s\right)\right]\times X_t^i\right\} \tag{27}$$

$$= E_T\left\{\sum_{s=T}^{+\infty}\sum_{t=T}^{s}\beta^{s-T}\gamma^{s-t}\frac{dh_s}{dZ_s}\times X_t^i\right\} \tag{28}$$

Similarly, the financial gain from defaulting on carbon debt at date T writes as follows:

$$E_T\left[\sum_{t=T}^{+\infty}\beta^{t-T}\mu_t Z_t^i\right] = E_T\left[\sum_{t=T}^{+\infty}\beta^{t-T}\frac{dh_t}{dZ_t}\times\left(\gamma^t Z_0^i+\sum_{s=1}^{t}X_s^i\gamma^{t-s}\right)\right] \tag{29}$$

$$= E_T\left[\sum_{t=T}^{+\infty}\beta^{t-T}\frac{dh_t}{dZ_t}\times\left(\gamma^{t-T}\left(\gamma^T Z_0^i+\sum_{s=1}^{T-1}X_s^i\gamma^{T-s}\right)+\sum_{s=T}^{t}X_s^i\gamma^{t-s}\right)\right] \tag{30}$$

$$= E_T\left[\sum_{t=T}^{+\infty}\beta^{t-T}\frac{dh_t}{dZ_t}\times\left(\gamma^{t-T}\times\gamma\times Z_{T-1}^i+\sum_{s=T}^{t}X_s^i\gamma^{t-s}\right)\right] \tag{31}$$

$$= E_T\left[\sum_{t=T}^{+\infty}\beta^{t-T}\frac{dh_t}{dZ_t}\times\left(\sum_{s=T}^{t}X_s^i\gamma^{t-s}\right)\right]+E_T\left[\sum_{t=T}^{+\infty}\left(\beta\gamma\right)^{t-T}\frac{dh_t}{dZ_t}\gamma Z_{T-1}^i\right] \tag{32}$$

$$= E_T\left[\sum_{t=T}^{+\infty}\sum_{s=T}^{t}\beta^{t-T}\gamma^{t-s}X_s^i\frac{dh_t}{dZ_t}\right]+E_T\left[\sum_{t=T}^{+\infty}\left(\beta\gamma\right)^{t-T}\frac{dh_t}{dZ_t}\right]\gamma Z_{T-1}^i \tag{33}$$

Comparing expressions (28) and (33), we obtain that defaulting on carbon debt is more tempting (i.e., less costly) than reneging on the carbon tax:

$$\Delta_{tax,T} - \Delta_{liability,T} = E_T\left[\sum_{t=T}^{+\infty}\beta^{t-T}\mu_t Z_t^i\right] - E_T\left[\sum_{t=T}^{+\infty}\beta^{t-T}\tau_t X_t^i\right] \tag{34}$$

$$= E_T\left[\sum_{t=T}^{+\infty}(\beta\gamma)^{t-T}\frac{dh_t}{dZ_t}\right]\gamma Z_{T-1}^i \tag{35}$$

$$= \tau_T\left(\gamma Z_{T-1}^i\right) > 0. \tag{36}$$

Notice that the magnitude of the difference is equal to the cost of the carbon stock inherited from the past, $\left(\gamma Z_{T-1}^i\right)$, priced at the carbon tax rate of the current period, τ_T. This difference does not account for the above-mentioned costs associated with defaulting on (financialized) carbon debt. These costs are relevant for the liability scheme but do not arise in the case of the carbon tax. Hence, the above difference constitutes an overestimate of the relative temptation to default when adopting a liability scheme rather than relying on a carbon tax.

A.2.2 On the relative temptation to delay the policy

On the other hand, a drawback of the carbon tax is that its adoption is costly up front. This is because it requires payments immediately for climate damage that may take decades or more to materialize. By contrast, a liability scheme asks for payments upon the realization of damage, effectively spreading payments over time. This makes the liability scheme more likely to be adopted in the short run. To be precise, compare the net benefits of both schemes over the first L periods, computed from the point of view of date zero:

$$\Delta_{liability-tax,L} = E_0\left\{\sum_{t=0}^{L-1}\beta^t\left[B_t^i\left(X_t^i\right) - \mu_t Z_t^i - D_t^i\left(Z_t\right)\right] - \sum_{t=0}^{L-1}\beta^t\left[B_t^i\left(X_t^i\right) - \tau_t X_t^i - D_t^i\left(Z_t\right)\right]\right\} \tag{37}$$

$$= E_0\left[\sum_{t=0}^{L-1}\beta^t\tau_t X_t^i\right] - E_0\left[\sum_{t=0}^{L-1}\beta^t\mu_t Z_t^i\right] \tag{38}$$

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

Computing each term separately, we obtain:

$$E_0\left[\sum_{t=0}^{L-1}\beta^t\tau_t X_t^i\right] = E_0\left[\sum_{t=0}^{L-1}\beta^t\frac{d}{dX_t^i}\left\{E_t\left[\sum_{s=t}^{+\infty}\beta^{s-t}h_s\left(Z_s\right)\right]\right\}X_t^i\right] \tag{39}$$

$$= E_0\left[\sum_{t=0}^{L-1}\beta^t\left\{\sum_{s=t}^{+\infty}\left(\beta\gamma\right)^{s-t}\frac{dh_s}{dZ_s}\right\}X_t^i\right] \tag{40}$$

$$= E_0\left[\sum_{t=0}^{L-1}\sum_{s=t}^{+\infty}\beta^s\gamma^{s-t}\frac{dh_s}{dZ_s}X_t^i\right] \tag{41}$$

$$= E_0\left[\sum_{s=0}^{+\infty}\sum_{t=0}^{\min\{s,L-1\}}\beta^s\gamma^{s-t}\frac{dh_s}{dZ_s}X_t^i\right] \tag{42}$$

$$= E_0\left[\sum_{s=0}^{L-2}\sum_{t=0}^{s}\beta^s\gamma^{s-t}\frac{dh_s}{dZ_s}X_t^i\right] + E_0\left[\sum_{s=L-1}^{+\infty}\sum_{t=0}^{L-1}\beta^s\gamma^{s-t}\frac{dh_s}{dZ_s}X_t^i\right] \tag{43}$$

$$= E_0\left[\sum_{s=0}^{L-2}\sum_{t=0}^{s}\beta^s\gamma^{s-t}\frac{dh_s}{dZ_s}X_t^i\right] + E_0\left[\beta^{L-1}\sum_{s=L-1}^{+\infty}\left(\beta\gamma\right)^{s-L+1}\frac{dh_s}{dZ_s}\sum_{t=0}^{L-1}\gamma^{L-1-t}X_t^i\right] \tag{44}$$

and

$$E_0\left[\sum_{t=0}^{L-1}\beta^t\mu_t Z_t^i\right] = E_0\left[\sum_{t=0}^{L-1}\beta^t\frac{dh_t}{dZ_t}\times\left(\sum_{s=0}^{t}X_s^i\gamma^{t-s}\right)\right] \tag{45}$$

$$= E_0\left[\sum_{t=0}^{L-1}\left(\beta\gamma\right)^t\frac{dh_t}{dZ_t}\times\left(\sum_{s=0}^{t}X_s^i\gamma^{-s}\right)\right] \tag{46}$$

$$= E_0\left[\sum_{t=0}^{L-1}\sum_{s=0}^{t}\beta^t\gamma^{t-s}\frac{dh_t}{dZ_t}X_s^i\right] \tag{47}$$

$$= E_0\left[\sum_{s=0}^{L-1}\sum_{t=0}^{s}\beta^s\gamma^{s-t}\frac{dh_s}{dZ_s}X_t^i\right] \tag{48}$$

$$= E_0\left[\sum_{s=0}^{L-2}\sum_{t=0}^{s}\beta^s\gamma^{s-t}\frac{dh_s}{dZ_s}X_t^i\right] + E_0\left[\sum_{t=0}^{L-1}\beta^{L-1}\gamma^{L-1-t}\frac{dh_{L-1}}{dZ_{L-1}}X_t^i\right] \tag{49}$$

Taking the difference, we get:

$$\Delta_{liability-tax,L} = E_0\left[\beta^{L-1}\sum_{s=L-1}^{+\infty}(\beta\gamma)^{s-L+1}\frac{dh_s}{dZ_s}\sum_{t=0}^{L-1}\gamma^{L-1-t}X_t^i\right] - E_0\left[\sum_{t=0}^{L-1}\beta^{L-1}\gamma^{L-1-t}\frac{dh_{L-1}}{dZ_{L-1}}X_t^i\right] \tag{50}$$

$$= \beta^{L-1}E_0\left[\sum_{s=L-1}^{+\infty}(\beta\gamma)^{s-L+1}\frac{dh_s}{dZ_s}\sum_{t=0}^{L-1}\gamma^{L-1-t}X_t^i\right] - \beta^{L-1}E_0\left[\frac{dh_{L-1}}{dZ_{L-1}}\sum_{t=0}^{L-1}\gamma^{L-1-t}X_t^i\right] \tag{51}$$

$$= \beta^{L-1}E_0\left[\sum_{s=L}^{+\infty}(\beta\gamma)^{s-L+1}\frac{dh_s}{dZ_s}\sum_{t=0}^{L-1}\gamma^{L-1-t}X_t^i\right] \tag{52}$$

$$= \beta^{L-1}E_0\left[(\beta\gamma)\,\tau_L Z_{L-1}^i\right], \tag{53}$$

so that delaying for L periods from the point of view of date zero writes:

$$\Delta_{liability-tax,L} = \beta^L E_0\left[\tau_L\left(\gamma Z_{L-1}^i\right)\right] > 0 \tag{54}$$

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.

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 $\tau_L = E_L\left[\sum_{s=L}^{+\infty}(\gamma\beta)^{s-L}\frac{\partial D_s}{\partial Z_s}\right]$ increases with L.[19] Therefore, the size of the difference can even increase with L if $\tau_{L+1}/\tau_L > 1/\beta$. With discount factors close to one, this is a distinct possibility.

A.2.3 Comparison

Evaluated from the point of view of date zero, the expected relative temptation to default on carbon debt writes:

$$\beta^T E_0\left[\Delta_{tax,T} - \Delta_{liability,T}\right] = \beta^T E_0\left[\tau_T\left(\gamma Z_{T-1}^i\right)\right] > 0 \tag{55}$$

Comparing with Expression (54), we see that the relative temptation to default on carbon debt mirrors the relative temptation to delay the implementation of

[19] As mentioned, the fact that the social cost of carbon increases over time is supported by most integrated assessment models.

a carbon tax.