Optimal Climate Policy

for a Pessimistic Social Planner

June 23, 2017

Abstract

This paper characterizes the preferences of a pessimistic social planner concerned with the potential costs of extreme, low-probability climate events. This pessimistic attitude is represented by a recursive optimization criterion *à la* [**?**] implying that a very sharp and early mitigation effort arises as the optimal climate policy. We find that for sufficiently high levels of risk-aversion an aggressive mitigation policy is chosen even when the discount factor is low. The dynamics of the optimal mitigation policy displays an *inverted policy ramp* with a sharp and immediate mitigation effort, followed by a gradual reduction until the pollution stock converges towards its long-run equilibrium. We also observe that the initial sharpness of the mitigation effort requires substantial capture and sequestration of carbon from the atmosphere. We extend our analysis showing that when the social planner observes the concentration and emission levels with a time lag, she undertakes a more aggressive policy to reduce the greater degree of uncertainty she faces. Finally, we show under which conditions the optimal mitigation policy dictated by our analysis coincides with that derived using the *robustness* approach of Hansen and Sargent (2008).

JEL numbers: C61, Q54.

Keywords: climate change, climate policy targets, risk-aversion, pessimism and precautionary principle, Linear Exponential Quadratic Gaussian.

1 Introduction

Climate change represents the most important real-life example of global externality and there is a wide consensus among economists on the need for mitigation effort to deal with it. However, economic models disagree on how sharp and how urgent such effort should be. A key issue in shaping climate policy is the impact of uncertainty. Such uncertainty is mainly due to the difficulty in forecasting climate change and its consequences and it is particularly relevant when we deal with the potential effects of extreme and catastrophic climatic events due to the reaching of some tipping point.

The uncertainty on where such tipping points are and what happens if we exceed them may justify the adoption of the precautionary principle so that a substantial effort should be made to avoid that extreme events occur. Indeed, Article 3.3 of the UN Framework Convention on Climate Change (UNFCCC) states: *"The Parties should take precautionary measures to anticipate, prevent or minimize the causes of climate change and mitigate its adverse effects. Where there are threats of serious or irreversible damage, lack of full scientific certainty should not be used as a reason for postponing such measures"*.

This paper addresses these issues by deriving the optimal climate policy of a risk-averse social planner whose preferences reflect her great concern about the consequences of extreme climate events. Our social planner conforms her decisions to the precautionary principle since she acts as if she were pessimistic and considered the worst-case environmental scenarios very likely.

Our findings partly confirm previous results as in the steady state of our model the optimal climate policy implies that the larger the degree of risk-aversion and/or the discount factor, the more aggressive the mitigation effort. However, risk-aversion seems to play a particularly important role. Indeed we see that for a sufficiently high level of risk-aversion a very aggressive mitigation policy will be chosen even when the discount factor is low.

When we look at the dynamics of the optimal climate policy, we observe an *inverted policy ramp* with a sharp and immediate mitigation effort, followed by a gradual reduction of such effort until the pollution stock converges to the pre-industrial level. This results depends on the social planner's preferences which entail that the optimal concentration level is equal to the pre-industrial one. We also observe that the higher the risk-enhancement coefficient, the earlier and more aggressive is the mitigation effort, and the faster is the convergence of pollution to its optimal level. Moreover, before converging to the long run equilibrium the emissions brought about by the optimal mitigation effort turn out to be negative. This is due to the presence in our model of strong inertia of GHGs in the atmosphere and the absence of natural absorbtion of emissions of greenhouse gases. Consequently the optimal climate policy calls for a substantial amount of capture and sequestration of carbon from the atmosphere which is necessary to obtain the socially optimal level in a reasonable spell of time.

Another result of our analysis is that for sufficiently high levels of risk-aversion there is no available mitigating policy which solves the social planner optimization exercise. However, when the social planner is not too risk-averse an optimal mitigation policy exists, it is unique and it is described by an explicit linear function of the emission and concentration levels.

Our results are achieved by formulating a Markovian discounted linear exponential quadratic Gaussian (DLEQG) problem that: *a*) involves a risk-enhancement coefficient that injects extra-convexity in the social planner's objective function *vis-a-vis* that of standard discounted linear-quadratic (DLQG) optimization problems; and *b*) allows to identify the optimal policy via a worst-case choice mechanism according to which the welfare loss is minimized against the most unfavorable event.

While in our base formulation the uncertainty the social planner faces is due to stochastic shocks which perturb the dynamics of the emission and concentration levels of greenhouse gases (GHGs), our analysis is also extended to study two more crucial dimensions of uncertainty.

Firstly, we introduce extra uncertainty into our base model by assuming that, not only are the emission and concentration levels of GHGs subject to stochastic shocks, but also that the social planner observes them either with a time lag or through a noisy signal. Under this extension the characteristics of the optimal mitigation policy do not change qualitatively, confirming our results in a more general formulation. Indeed, we observe that the social planner will undertake a more aggressive mitigating policy in order to reduce the greater degree of uncertainty she faces.

Secondly, we consider the case in which the social planner is concerned with the possibility that her assumptions on the dynamics of the emission and concentration levels of GHGs may be incorrect. Interestingly, given an uncertain law of motion and assuming a null risk-enhancement coefficient, the optimal mitigating policy chosen by the social planner according to [?] *robustness* approach coincides with that which applies to our base DLEQG formulation.

Therefore, our model can be interpreted in two alternative ways, as either representing a risk-averse social planner or one uncertain about the dynamics of emission and concentration levels. In the latter case, our analysis is related to that of [?], who employ the Hansen and Sargent's robustness framework to a similar problem of optimal pollution control. However, in their formulation the social planner is not allowed to learn over time about the mis-specification of her assumptions on the dynamics of the state variables. This weakness of Hansen and Sargent's framework is absent in the DLEQG formulation we consider as no mis-specification of the dynamics of the state variables is introduced. In fact, differently from [?], we allow for the possibility that the social planner observes imperfectly the emission and concentration levels of GHGs and learns over time such values through the information contained in some noisy signal. Moreover, with respect to [?] we consider a discrete-time model which can accommodate the case of a social planner facing a terminal horizon of intervention.

More generally, both our paper and [?] contribute to the literature studying the effects of risk and uncertainty in frameworks characterized by climate and economic dynamics (see [?], [?] and the more recent contributions by [?], [?], [?], [?], [?], [?], [?] and [?], among others). In particular, the possibility of separating risk-aversion from intertemporal preferences makes our paper close to [?], who discusses how this disentanglement applies to a model with Epstein-Zin's preferences and uses this argument to justify the low discount rate chosen by Stern (2007).[1]

In addition to the risk-sensitive characterization of the social planner preferences, we contribute to the extant literature in other ways. Firstly, differently from existing models, in our base formulation the emission and concentration levels are both subject to stochastic shocks, so that the social planner does not have full control on emission flows, and a costly effort is required to reduce emissions. Secondly, we propose a different characterization of social costs *vis-a-vis* other related papers (see, for instance, [?], [?] and [?]) as we assume that while the reduction in carbon emissions requires such costly effort it does not impair output levels.

Albeit new in the literature we think that this assumption is compatible with some new views about sustainable growth. In fact, the common assumption of a positive relation between emissions and growth has been recently challenged by some leading economists (Philippe

[1]We refer to [?] for an in-depth analysis of how risk-sensitiveness may actually act through the discount rate. On the the separation between risk-aversion and intertemporal preferences see also [?], [?] and [?].

Aghion, Daniel Kahneman, Ian Parry, Torsten Persson, Michael Spence and Nicholas Stern, among others), policymakers and opinion makers from the business and finance community who contributed to the New Climate Economy Report, 2014. This report is part of a more general project which is aimed at collecting evidences supporting the thesis that climate policy is now compatible with strong economic growth.[2] Further empirical support for this thesis has been provided by [**?**]. They identify the effect of the Climate Change Levy on manufacturing imposed by the UK government in 2001, showing that this policy had a strong negative impact on energy intensity, and therefore on carbon emissions, but it did not impair employment, revenues, TFP and investment. Assuming independence between climate policy and output levels does not mean that we do not recognize the costs of mitigation. In our formulation the cost of the effort required to reduce emissions and that due to the concentration of GHGs in the atmosphere are both quadratic and hence potentially very large. Such costs coupled with our definition of the social planner's preferences allow to identify an aggressive optimal mitigation policy which entails rapid convergence to the pre-industrial level of GHG's concentration.

The paper is organized as follows. In the next Section we describe the analytical formulation of our model and derive the optimal mitigation effort in steady state. In Section 3 we analyze the properties of the optimal mitigation policy by performing a numerical analysis. Section 4 and Section 5 are dedicated to extensions of our base formulation. In the former we investigate the impact of the imperfect observation of the emission and concentration levels on the part of the social planner. In the latter we discuss the link between our risk-sensitive formulation and the *robustness* approach *à la* Hansen and Sargent. A final Section provides concluding remarks. Proofs and other analytical results are relegated in Appendix A, while Appendix B contains an explanation of how we have calibrated the parameters employed in the numerical analysis.

2 The Model

We define a discrete-time stochastic dynamic model for the stock and emission levels of GHGs in the atmosphere. In this model p_t denotes the deviation of the concentration level of GHGs at time t from its pre-industrial level while e_t denotes the level of anthropogenic emission flow of GHGs in the interval $(t-1, t]$. Uncertainty is introduced in the model through two

[2]See for instance the two country cases on USA "Creating a New Climate Economy in the United States" and China "China and The New Climate Economy" which are available at http://newclimateeconomy.report/.

idiosyncratic terms, ϵ_t^p and ϵ_t^e, that affect the dynamics respectively of p_t and e_t and a third idiosyncratic term, η_t, representing a random component of the cost function associated with both pollution and the mitigation effort. The introduction of the stochastic term ϵ_t^p can be justified by the impossibility to define exactly the atmospheric lifetime of GHGs. For example, CO2 has a variable atmospheric lifetime which is estimated in a range between 30 and 95 years (see Archer et al., 2009) and other greenhouse gases show similar features. Moreover, we introduce the stochastic element ϵ_t^e into the dynamics of e_t because society does not possess perfect control over the emission level, while η_t captures uncertainty on the pollution costs associated with the concentration of GHGs because of the indeterminacy in the impact of such concentration on global warming and the environment.

Without any intervention to curb emission the dynamics of GHGs concentration is

$$p_{t+1} \;=\; \gamma\, p_t \,+\, e_{t+1} \,+\, \epsilon_{t+1}^p\,, \tag{2.1}$$

where $\gamma \in [0,1]$ is a constant term capturing the persistence of the stock of GHGs and ϵ_t^p is a white noise process with $\epsilon_t^p \sim N(0,\sigma_p^2)$. The higher γ, the lower the environment's absorptive capacity with respect to a specific pollutant. Indeed, if $\gamma = 0$ the pollutant will produce its effects only in the period it has been emitted (*flow* pollutant). On the contrary, if $\gamma = 1$ the pollutant is maximally persistent since the environment has no absorptive capacity (*stock* pollutant). GHGs are generally defined as pollutants for which the environment has some absorptive capacity (i.e. $0 < \gamma < 1$).[3]

Let $\Delta e_{t+1} = e_{t+1} - e_t$ denote the variation in the emission level across periods. If this value is negative, we observe abatement in GHGs emissions between period t and $t+1$. Δe_{t+1} depends on the control variable u_t representing the effort exerted at time t by society to reduce the impact of human activity on the environment. We assume that the effort u_t is selected once the concentration level, p_t, and the emission level, e_t, have been observed. Then,

$$e_{t+1} \;=\; e_t \,+\, u_t \,+\, \epsilon_{t+1}^e, \tag{2.2}$$

where ϵ_t^e is a white noise process, with $\epsilon_t^e \sim N(0,\sigma_e^2)$, independent of ϵ_t^p. In the absence of the shock ϵ_{t+1}^e a reduction in pollution is possible only for $u_t < 0$, so that a mitigation effort

[3] We abstract from the possibility that also e_t can be partly absorbed by the atmosphere. Dealing with this possibility would imply the introduction of an additional parameter that, however, would not qualitatively affect our findings.

corresponds to a negative control. Note that (2.2) allows for the possibility of sequestrating emissions since e_{t+1} can be negative for sufficiently high values of the mitigation effort. Hence, substituting (2.2) in (2.1), we conclude that

$$p_{t+1} \;=\; \gamma\, p_t \,+\, e_t \,+\, u_t \,+\, \epsilon^e_{t+1} \,+\, \epsilon^p_{t+1}\,. \tag{2.3}$$

We can regroup equations (2.2) and (2.3) in a Markovian formulation represented by the following law of motion for the state vector $\mathbf{z}_t$

$$\mathbf{z}_{t+1} \;=\; \mathbf{A}\,\mathbf{z}_t \,+\, \mathbf{B}\,u_t \,+\, \boldsymbol{\epsilon}_{t+1}\,, \tag{2.4}$$

where

$$\mathbf{z}_t \equiv \begin{pmatrix} p_t \\ e_t \end{pmatrix}, \quad \boldsymbol{\epsilon}_t \equiv \begin{pmatrix} \epsilon^e_t + \epsilon^p_t \\ \epsilon^e_t \end{pmatrix}, \quad \mathbf{A} \equiv \begin{pmatrix} \gamma & 1 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{B} \equiv \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and $\boldsymbol{\epsilon}_t \sim N(\mathbf{0},\mathbf{N})$, with $\mathbf{N} \equiv \begin{pmatrix} \sigma^2_p + \sigma^2_e & \sigma^2_e \\ \sigma^2_e & \sigma^2_e \end{pmatrix}$ the covariance matrix of the shock terms in equations (2.2) and (2.3).

We assume that the costs associated with pollution and effort are both quadratic. This assumption reflects the evidence that both marginal environmental damages and marginal abatement costs are positive and increasing respectively in pollution and the abatement effort. Specifically, the per period cost function is

$$c_t \;=\; \beta\, p^2_t \,+\, \alpha\, u^2_t \,+\, \eta_t\,, \tag{2.5}$$

where α and β are two positive constants and η_t is a white noise process, with $\eta_t \sim N(0,\sigma^2_\eta)$ independent of both ϵ^e_t and ϵ^p_t. In our formulation the environmental damages are minimized when $p_t = 0$. This implies that we are normalizing to zero the pre-industrial or natural level of GHGs' stock.

This formulation abstracts from the potential benefits of pollution. These benefits are usually characterized in terms of output (see, for instance, [**?**] and [**?**]), so that the cost of reducing GHGs are identified with the supposed lower economic activity that climate policies would bring about. While we recognize that society must afford a costly effort to reduce emissions, we do not consider in our formulation any causal link between the emissions of

GHGs and output levels. In fact, we do not see a larger output level and the reduction of emissions of GHGs as conflicting goals. On the contrary, we share the view of the recent New Climate Economy Report [?], according to which raising resource efficiency, investing in infrastructure and stimulating innovation in three key economic areas - namely urban policies, land use and energy - *"can make growth and climate objectives mutually reinforcing in both the short and medium term. In the long term, if climate change is not tackled, growth itself will be at risk"* [? , p. 7].[4] Consequently, in our formulation costs are only associated with the effort to reduce the emissions of GHGs and the impact of their concentration on the environment.

For analytical convenience we re-write the cost function (2.5) in matrix form, by introducing the matrices

$$\mathbf{Q} \equiv \alpha\,, \quad \mathbf{R} \equiv \begin{pmatrix} \beta & 0 \\ 0 & 0 \end{pmatrix}\,, \text{ so that}$$

$$c_t \quad = \quad c_{o,t} \;+\; \eta_t\,, \quad \text{where} \quad c_{o,t} \;=\; \mathbf{Q}\,u_t^2 \;+\; \mathbf{z}_t'\,\mathbf{R}\,\mathbf{z}_t\,. \tag{2.6}$$

To represent the social planner's preferences capturing pessimism against the realizations of extreme events, we employ a discounted linear-exponential-quadratic Gaussian (DLEQG) recursive model *à la* [?] as formulated in [?]. The DLEQG characterization of the social planner's preferences involves the presence of a risk-enhancement coefficient that injects extra convexity in her objective function *vis-a-vis* that of a standard discounted linear-quadratic Gaussian (DLQG) formulation. This extra convexity is crucial in our analysis since it captures the importance of extreme, possibly catastrophic outcomes in the decisions of the policy makers. More specifically, we assume that the social planner solves the following recursive optimization

$$\exp\left(\frac{\rho}{2}\,\boldsymbol{\mathcal{V}}_t\right) \quad = \quad \min_{u_t}\, E_t\left[\exp\left(\frac{\rho}{2}(c_t \;+\; \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right]\,, \tag{2.7}$$

where ρ (with $\rho > 0$) is the *risk-enhancement* coefficient, δ (with $0 < \delta < 1$) is the time-discounting coefficient and $\boldsymbol{\mathcal{V}}_t$ is the value function over the periods $t = 1, 2, \ldots, \infty$ with respect

[4] On the same issue see also [?] who show that under plausible conditions an optimal climate policy accomplishing a transition towards clean technologies would not sacrifice long-run growth. As already mentioned before, [?] find that the carbon tax imposed on manufacturing plants under the Climate Change Levy package - the most important climate policy that the UK government has introduced so far - did not affect their productivity and competitiveness. In 2016 the decoupling of global emissions and economic growth has been reported also by the International Energy Agency in its latest release of data on OECD energy and emission trends [?].

to the free-valued control u_t.[5] To appreciate the role of the *risk-enhancement* coefficient, ρ, and provide an economic interpretation for the value function $\boldsymbol{\mathcal{V}}_t$, we observe that when $\rho \downarrow 0$ solving the recursive optimization in (2.7) is equivalent to solving the Bellman equation for a discounted linear quadratic Gaussian (DLQG) problem (see Appendix A.2),

$$\boldsymbol{\mathcal{V}}_t \;=\; \min_{u_t} \left\{ c_{o,t} \,+\, \delta\, E_t\, [\boldsymbol{\mathcal{V}}_{t+1}] \right\} .$$

In this limit case the value function corresponds to the expected discounted value of all future costs associated with pollution and the abatement effort, $\boldsymbol{\mathcal{V}}_t = E_t \left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} c_{o,\tau} \right]$. When $\rho > 0$ the value function is no longer time-additive, but it still represents an aggregate measure of the social costs associated with pollution and abatement.[6] Importantly, in this new formulation large values for the pollution levels and abatement efforts are associated with larger social costs than those met for $\rho \downarrow 0$, so that by introducing ρ into the social planner's preferences we increase her degree of risk-aversion.[7]

To pin down the optimal mitigation policy which prevails when the law of motion for the state vector is defined as in equation (2.3) and the cost function is as in equation (2.6), we introduce the function $\boldsymbol{\mathcal{S}}_t \equiv [c_{o,t} + \delta \boldsymbol{\mathcal{V}}_{t+1} - \frac{1}{\rho} \boldsymbol{\epsilon}'_{t+1} \mathbf{N}^{-1} \boldsymbol{\epsilon}_{t+1}]$. The function $\boldsymbol{\mathcal{S}}_t$ is a combination of the current cost function $c_{o,t}$, the future value function $\boldsymbol{\mathcal{V}}_{t+1}$, which represents an aggregate measure of the future social costs associated with pollution and the abatement effort, and the penalization $-\frac{1}{\rho} \boldsymbol{\epsilon}'_{t+1} \mathbf{N}^{-1} \boldsymbol{\epsilon}_{t+1}$ for the uncertainty over future values of the state variables, p_{t+1} and e_{t+1}.

Then, to identify the optimal effort which prevails in the steady state we rely on the following Lemma, whose proof based on results in [**?**] is discussed in Appendix A.4.

Lemma 1 *If the value function $\boldsymbol{\mathcal{V}}_{t+1}$ is a quadratic form in $\mathbf{z}_{t+1}$ and the function $\boldsymbol{\mathcal{S}}_t$ satisfies a saddle point condition with respect to $\boldsymbol{\epsilon}_{t+1}$ and u_t, so that $\min_{u_t} \max_{\boldsymbol{\epsilon}_{t+1}} \boldsymbol{\mathcal{S}}_t$ exists, then the*

[5]Appendix A.1 shows that in solving (2.7) the idiosyncratic shock η_t does not play any role, so that we can substitute c_t with $c_{o,t}$ in the recursive optimization problem.

[6]Exploiting results from [**?**] and [**?**] one can prove that the recursive criterion in (2.7) corresponds to the [**? ?**] recursive utility where the log of the consumption level is a quadratic function of the state vector, $\mathbf{z}_t$, and the control value, u_t, and the inter-temporal rate of substitution is 1 (see Appendix A.3).

[7]This results holds because it can be proved (see [**?**]) that $\boldsymbol{\mathcal{V}}_t$ is a convex function of the state vector $\mathbf{z}_t$ and that such convexity augments with ρ.

following equality holds

$$\min_{u_t} E_t \left[\exp\left(\frac{\rho}{2}(c_{o,t} \; + \; \delta\, \boldsymbol{\mathcal{V}}_{t+1})\right)\right] \quad = \quad \exp\left(\frac{\rho}{2}\left[\nu_t + \min_{u_t} \max_{\boldsymbol{\epsilon}_{t+1}} \boldsymbol{\mathcal{S}}_t\right]\right),$$

where ν_t *is a constant independent of* $\mathbf{z}_t$*, while the value function* $\boldsymbol{\mathcal{V}}_t$ *is a quadratic form in* $\mathbf{z}_t$ *equal to the extremized function* $\boldsymbol{\mathcal{S}}_t$*,* $\min_{u_t} \max_{\boldsymbol{\epsilon}_{t+1}} \boldsymbol{\mathcal{S}}_t$*, plus a constant independent of* $\mathbf{z}_t$*.*

This result means that an optimum is reached when the function $\boldsymbol{\mathcal{S}}_t$ satisfies a *saddle point condition*, according to which first $\boldsymbol{\mathcal{S}}_t$ is maximized with respect to $\boldsymbol{\epsilon}_{t+1}$ and then the resulting function is minimized with respect to u_t.

An economic interpretation of this condition is that a risk-averse social planner whose preferences are represented by the optimization criterion (2.7) attempts to hedge against the worst possible values for the vector $\boldsymbol{\epsilon}_{t+1}$, by following a *min-max* strategy according to which she selects u_t to minimize the function $\boldsymbol{\mathcal{S}}_t$ against the most unfavorable vector $\boldsymbol{\epsilon}_{t+1}$. Such a social planner acts *as if* she were pessimistic, considering these *worst-case* realizations very likely. Consequently the social planner tunes her actions on their impact on the function $\boldsymbol{\mathcal{S}}_t$, applying what we term, borrowing Whittle's terminology [**?**], a *pessimistic* (or worst-case) choice mechanism.

Exploiting Lemma 1 we can then characterize the optimal effort which prevails in the steady state. In particular, we can establish that if the social planner is not too pessimistic, that is if ρ is not too large, an optimal mitigation policy exists and it is given by a linear function of the emission and concentration levels, in that two coefficients κ_p and κ_e exist such that in steady state the optimal mitigation effort in t is

$$u_t \;=\; \kappa_p p_t \;+\; \kappa_e e_t\,.$$

The conditions under which such equilibrium exists and the values taken by the two coefficients κ_p and κ_e are spelled out in the following Proposition.[8]

Proposition 1 *If* $\mathbf{\Pi}$ *is a (2 by 2) semi-positive definite symmetric matrix such that* $(\delta\mathbf{\Pi})^{-1}-$

[8] We only investigate the characteristics of linear equilibria. In a non-linear equilibrium the optimal mitigation policy is some unspecified non-linear function of the state variables. The analysis of this class of equilibria is beyond the scope of our contribution.

$\rho\mathbf{N}$ *is positive definite and it represents a fixed point in the following equation*

$$\mathbf{\Pi} \;=\; \mathbf{R} + \mathbf{A}'\widetilde{\mathbf{\Pi}}\mathbf{A} - \mathbf{A}'\widetilde{\mathbf{\Pi}}\mathbf{B}(\mathbf{Q} + \mathbf{B}'\widetilde{\mathbf{\Pi}}\mathbf{B})^{-1}\mathbf{B}'\widetilde{\mathbf{\Pi}}\mathbf{A}\,, \;\textit{with}\; \widetilde{\mathbf{\Pi}} \;=\; ((\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N})^{-1} \tag{2.8}$$

in steady state the optimal mitigation effort in t is linear in the concentration and emission levels, $u_t = \kappa_p p_t + \kappa_e e_t$, *with*

$$\kappa_p \;=\; -\gamma\left(1 - \frac{\alpha + \tilde{\pi}_{1,2} + \tilde{\pi}_2}{\alpha + \tilde{\pi}_1 + 2\,\tilde{\pi}_{1,2} + \tilde{\pi}_2}\right), \tag{2.9}$$

$$\kappa_e \;=\; -\left(1 - \frac{\alpha}{\alpha + \tilde{\pi}_1 + 2\,\tilde{\pi}_{1,2} + \tilde{\pi}_2}\right) \tag{2.10}$$

and $\tilde{\pi}_1$, $\tilde{\pi}_2$ *and* $\tilde{\pi}_{1,2}$ *the elements of the matrix* $\widetilde{\mathbf{\Pi}}$, *while the value function is* $\boldsymbol{\mathcal{V}}_t = \mathbf{z}_t'\mathbf{\Pi}\mathbf{z}_t + \lambda$ *with*

$$\lambda \;=\; \frac{1}{1-\delta}\left(\frac{1}{4}\rho\sigma_\eta^2 - \frac{1}{\rho}\ln(det[\mathbf{I} - \delta\rho\mathbf{N}\mathbf{\Pi}])\right).$$

Proposition 1 posits that to find the optimal mitigation policy in steady-state the modified Riccati equation (2.8) must be solved,[9] under the condition that the matrix $(\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N}$ being positive definite. This condition is required for the recursive optimization in (2.7) to have a meaningful solution. In fact, we have seen in Lemma 1 that such a solution exists if the function $\boldsymbol{\mathcal{S}}_t$ satisfies a saddle point condition. In [**?**] it is proved that for the function $\boldsymbol{\mathcal{S}}_t$ to possess a saddle point and the recursive optimization (2.7) to have a solution the matrix $(\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N}$ must be positive definite.

For a sufficiently large degree of risk-aversion (i.e. for a large enough ρ) $(\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N}$ will not be positive definite, indicating that such saddle point condition cannot be met and that the recursive optimization does not admit a solution, in that the value of $\boldsymbol{\mathcal{V}}_t$ becomes infinite. In other words, for a sufficiently large degree of pessimism, i.e. for ρ large enough, the optimization exercise we investigate is not well-behaved and hence it becomes meaningless. A possible economic interpretation of the failure of the optimization to have a proper solution is that in these extreme circumstances the social planner becomes so pessimistic as to consider her attempt to reduce the emission level ineffective and hence useless.

Proposition 1 indicates that when the condition that $(\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N}$ is positive definite is met

[9]For $\rho = 0$ equation (2.8) coincides with the Riccati equation which characterizes a standard DLQG problem.

and a steady-state equilibrium exists, the optimal mitigation policy is no longer independent of the degree of uncertainty of the social planner, as now u_t depends on the covariance matrix of the vector of shocks, $\mathbf{N}$. In fact, such matrix enters into the determination of the coefficients κ_p and κ_e in equations (2.9) and (2.10) through the matrix $\widetilde{\mathbf{\Pi}} = ((\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N})^{-1}$. Moreover, because of the nature of the cost function c_t, it can be established that $-\gamma < \kappa_p < 0$ and $-1 < \kappa_e < 0$. Therefore, the optimal policy requires some mitigation effort but it does not entail that the pollution level immediately converges to the pre-industrial one.

The values taken by the coefficients κ_p and κ_e depend on the matrix $\mathbf{\Pi}$. This matrix is a fixed point in a highly non-linear system of equations. This does not have an explicit solution and numerical methods are called for. A possible straightforward strategy is to guess the initial solution $\mathbf{\Pi} = \mathbf{0}$ and then apply the modified Riccati equation sequentially (which corresponds to a backward recursion) through the concatenation of $\widetilde{\mathbf{\Pi}} = ((\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N})^{-1}$ and equation (2.8) until convergence. This may fail. In addition, more than one fixed point could exist. However, using results from [**?**] we can establish that a sufficient condition for a unique steady state is that: i) $\mathbf{Q}$ is positive definite; and ii) for some r, $\sum_{m=0}^{r-1}(\mathbf{A}')^m\mathbf{R}(\mathbf{A})^m$ and $\sum_{m=0}^{r-1}(\sqrt{\delta}\mathbf{A}')^m\mathbf{J}(\sqrt{\delta}\mathbf{A})^m$ are positive definite, where $\mathbf{J} = (\sqrt{\delta}\mathbf{B})'\mathbf{Q}^{-1}(\sqrt{\delta}\mathbf{B}) - \delta\rho\mathbf{N}$. In Appendix A.6 we show that condition i) holds and that condition ii) holds for ρ small enough, so that we can confidently conjecture that the steady state, when it exists, is actually unique.

In performing the numerical procedure to seek out the steady state one must be careful in checking that $\mathbf{\Pi}$ is invertible. If this is not the case, Proposition 1 must be amended. In particular, the second order condition is now that $\delta\mathbf{\Pi} - \frac{1}{\rho}\mathbf{N}^{-1}$ is negative definite, while the risk-adjustment equation $\widetilde{\mathbf{\Pi}} = ((\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N})^{-1}$ becomes

$$\widetilde{\mathbf{\Pi}} \quad = \quad \delta\mathbf{\Pi}\,(\mathbf{I} - \delta\,\rho\mathbf{N}\mathbf{\Pi}\,)^{-1}\;. \tag{2.11}$$

From this we immediately see that when $\mathbf{\Pi} = \mathbf{0}$ then $\widetilde{\mathbf{\Pi}} = \mathbf{0}$ while the second order condition, that $\delta\mathbf{\Pi} - \frac{1}{\rho}\mathbf{N}^{-1}$ being negative definite, is trivially satisfied.

3 Optimal Mitigation Policy

In this Section we discuss the properties of the optimal mitigation policy. Because of the complexity of the closed-form characterization of the equilibrium in Proposition 1, we resort to numerical analysis to investigate its features. We analyze the dependence of the optimal mitigation policy on the key parameters of the model in steady state and its dynamics within a finite-horizon formulation. We have experimented with several alternative parametric configurations. We have consistently found that convergence of the numerical procedure presented in the previous Section is reached, within a short number of iterations, as long as ρ, the risk-enhancement coefficient, is not too large. In addition, while the results presented in this Section concern our base parametric choice, qualitatively similar findings arise in all alternative specifications we have investigated.[10]

The parameters of the baseline configuration have been carefully calibrated as follows: $\alpha = 30$, $\beta = 0.01115$, $\delta = 0.97$, $\gamma = 0.9917$ $\rho = 0.5$, $\sigma_e^2 = 0.01$ and $\sigma_p^2 = 0.0549$. These values, chosen considering a yearly time scale and 2012 as base year, are when possible coherent with corresponding ones selected by [**?**] and [**?**]. Their selection is presented in detail in Appendix B.

Albeit realistic, the numerical results presented in this Section should be assessed with some caution. Indeed, although we use parameters' values which are as close as possible to those employed in the previous literature, our model does not contain the details of integrated assessment models like DICE and the departure of our formulation from those employed in any other paper represents an obstacle to any quantitative comparison.

[Figure 1 about here.]

In Figure 1 we plot the dependence of the optimal policy coefficients, κ_p and κ_e, and the optimal abatement effort, u_t, given the concentration and emission levels in 2012, on the risk-enhancement coefficient, for ρ ranging from 0 to 1 and for three alternative values of the discount factor δ ($\delta = 0.95$, $\delta = 0.97$ and $\delta = 0.99$). Exploiting results from [**?**] we show in Appendix A.7 that, depending on δ, for ρ ranging between 0 and 1 the coefficient of relative risk-aversion of Epstein and Zin's recursive preferences varies from 1 to 50, an interval of values consistent with those usually employed in the economic literature. Moreover, for ρ between 0 and 1 the second order condition imposed by Proposition 1, that the matrix $(\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N}$ is positive definite, is respected.

The optimal mitigation effort plotted in the Figure is obtained applying the optimal control

[10] Such results are reported in a separate web Appendix.

rule, $u_t = \kappa_p p_t + \kappa_e e_t$, to the concentration and emission levels observed in 2012. In 2012 the emission level of CO2 was equal to 9.7GtC, while the corresponding concentration level was 836GtC (see, for instance, CO2Now.org). Considering that the pre-industrial concentration level was 590GtC, we conclude that in year 2012 e_t and p_t were equal respectively to 9.7 and 246. This implies that, given the values of κ_p and κ_e in the top and middle panels of Figure 1, the mitigation effort in 2012, u_t, ranges between -4.4GtC and -6.8GtC.

Figure 1 clearly indicates that the risk-enhancement coefficient ρ heavily influences the mitigation policy, as a larger ρ induces the social planner to act more aggressively (the coefficients κ_p and κ_e are larger in absolute value for ρ larger) and choose to reduce by a larger quantity ($-u_t$ is larger for ρ larger) the emission of GHGs, confirming that the extra convexity this risk-enhancement coefficient imposes on social preferences brings about a more aggressive policy. In particular, when $\delta = 0.97$, for $\rho = 0$ u_t is about -4.9GtC, while for $\rho = 0.5$ it is about -5.4GtC, with an increase in the abatement effort of about 10%.

We also see that, for any value of ρ, κ_p and κ_e and $-u_t$ are larger for $\delta = 0.99$ than for $\delta = 0.97$ and for $\delta = 0.97$ than for $\delta = 0.95$. In other words, the mitigation policy is more aggressive when the discount factor is larger. This confirms the intuition that the larger the weight attached to future costs, the more aggressively the social planner will reduce the emission level.

Importantly, Figure 1 also suggests that for a sufficiently high level of risk-aversion an aggressive policy will be chosen even when the discount factor is low. This implies that the policy recommendation of a sharp reduction in GHGs arises under both the value for the discount rate employed by [**?**] and that used by [**?**] and many others.

[Figure 2 about here.]

In Figure 2 we plot the trajectories of the concentration and emission levels, and of the optimal abatement effort, derived from a simulation of the equilibrium dynamics. Specifically, considering the starting values for the concentration and emission levels in 2012, the dynamics of p_t, e_t and u_t is simulated over 80 years according to the steady state equilibrium using randomly generated values for the shocks ϵ_t^p and ϵ_t^e for three values of ρ and for the baseline choice of the other parameters.

According to this simulated dynamics the pollution stock converges over the long-run to the pre-industrial level within 3/4 decades. This rapid convergence to the long-run stationary equilibrium is brought about by an aggressive mitigation policy which prescribes a sharp and early contraction in the emission level. Indeed, we observe an *inverted policy ramp* in that the social planner starts with a very aggressive mitigation policy and then reduces gradually her effort. Such mitigation effort requires a significant level of sequestration of GHGs at the beginning of the convergence process followed by an overshooting of the pre-industrial level before converging to it.

This particular dynamics is due to the combination of three features of our reduced form model: the social planner's preferences, which entail that the optimal concentration level is equal to the pre-industrial one; the strong inertia of GHGs in the atmosphere; and the absence of natural absorbtion of emissions of greenhouse gases. These features imply that in order to achieve convergence of the pollution concentration to the pre-industrial level an aggressive mitigation policy is called for. This requires a substantial amount of carbon capture and sequestration with a peak of about 13GtC by the end of the first decade. Such amount of capture and sequestration is less surprising than it may appear at first glance. Indeed, the sequestration levels displayed in the middle panel of Figure 2 are of the same order of magnitude of other simulations aimed at investigating emission pathways consistent with the achievement of the Paris 2015 target [**?**]. Such mitigation policy would not be feasible with the current technologies. However, the call for sequestration of GHGs from the atmosphere is now part of the debate on what is required to curb global warming and developing new *negative emission technologies* is one of the most urgent research challenges (Fuss et al., 2016).

We also observe that the presence of ρ does not affect qualitatively the dynamics of concentration and emission levels even if it has a quantitative impact. Indeed, the larger the risk-enhancement coefficient, the earlier and more aggressive is the mitigation effort and the faster is the convergence towards the steady state. When ρ is larger, so that the social planner attaches more weight to extreme events, she finds it optimal to exert a more pronounced effort in the first few years of her intervention in order to anticipate the achievement of her mitigation targets. In particular, as illustrated in the bottom panel of Figure 2, the initial abatement effort, u_t, with $\rho = 1$ is about 25 percent larger than that with $\rho = 0$. With a larger ρ the optimal mitigation policy brings about a faster reduction of the pollution levels in the first twenty years and a subsequent earlier convergence to the pre-industrial levels. This is shown in the top and middle panels of Figure 2 where we see that for $\rho = 1$ concentration and emission levels converge to pre-industrial levels within about 30 years while with $\rho = 0$ the same convergence is reached after nearly 50 years.

[Figure 3 about here.]

In Figure 3 we plot the dependence of the unconditional standard deviations $\sigma(p_t)$ and $\sigma(e_t)$ on ρ for the same alternative values of the discount factor ($\delta = 0.99$, $\delta = 0.97$ and $\delta = 0.95$) as in Figure 1. The unconditional variances $\text{Var}[p_t]$ and $\text{Var}[e_t]$ are components of the unconditional covariance matrix of the state vector $\mathbf{z}_t$. To obtain such unconditional covariance matrix consider that $\mathbf{z}_{t+1} = \mathbf{A}\mathbf{z}_t + \mathbf{B}u_t + \boldsymbol{\epsilon}_{t+1}$, where in equilibrium $u_t = \mathbf{K}\mathbf{z}_t$ (with $\mathbf{K} = (\kappa_p \, \kappa_e)$). This implies that $\mathbf{z}_{t+1} = \boldsymbol{\Gamma}\mathbf{z}_t + \boldsymbol{\epsilon}_{t+1}$, with $\boldsymbol{\Gamma} = \mathbf{A} + \mathbf{B}\mathbf{K}$, or equivalently $\mathbf{z}_t = \boldsymbol{\Lambda}\boldsymbol{\epsilon}_t$, where $\boldsymbol{\Lambda} = (\mathbf{I}_2 - \boldsymbol{\Gamma}L)^{-1}$ and L is the lag operator.

In Appendix A.8 we show that in steady state unconditionally $\mathbf{V} \equiv \text{Var}[\mathbf{z}_t] = \boldsymbol{\Lambda}\mathbf{N}\boldsymbol{\Lambda}'$ and, considering that $u_t = \mathbf{K}\mathbf{z}_t$, $\sigma_u^2 \equiv \text{Var}[u_t] = \mathbf{K}\,\mathbf{V}\,\mathbf{K}'$. Importantly, given the expressions for κ_p and κ_e, in Appendix A.8 we also show that $\mathbf{K}\boldsymbol{\Lambda} = (0 \;\; -1)$, so that $\sigma_u^2 = \sigma_e^2$. This indicates that in steady state the unconditional standard deviation of the reduction in emission, $\sigma(u_t)$, is independent of the degree of risk-aversion of the social planner, ρ. This means that empirically to discern among different levels of risk-aversion one needs to look at the volatility of the pollution level.

Figure 3 shows that $\sigma(p_t)$ and $\sigma(e_t)$ are smaller for $\delta = 0.99$ than for $\delta = 0.97$ and for $\delta = 0.97$ than for 0.95. The risk-enhancement coefficient ρ exerts a pronounced impact on the unconditional standard deviations $\sigma(e_t)$ and $\sigma(p_t)$. This indicates that a sufficiently high level of risk-aversion will result in a smaller volatility of the GHGs concentration level even when the discount factor is low, because a very aggressive mitigation policy will be chosen as we have seen in Figure 1. Such a relation is not surprising in that, given the extra convexity in the objective function of the recursive optimization in (2.7) brought about by a positive ρ, the social planner prefers early resolution of uncertainty and hence a smaller volatility of the state vector.[11]

[Figure 4 about here.]

[11] More precisely, as shown by [**?**], in the recursive optimization in (2.7) the elasticity of inter-temporal substitution is one. For $\rho > 0$ the objective function of the recursive optimization in (2.7) presents a coefficient of relative risk-aversion that is larger than one and hence it is greater than the inverse of the inter-temporal elasticity of substitution, the condition under which, according to [**?**], the social planner prefers early resolution of uncertainty. See also [**?**] and Appendix A.9.

In Figure 4 we plot the dynamics of the coefficients k_e and k_p of the optimal mitigation policy, alongside the abatement effort u_t, in the run up to the terminal date of the finite horizon formulation. We compare the dynamics of these values with the corresponding ones prevailing in the steady state formulation. We consider two scenarios: in the former the terminal date is 40 periods ahead, in the latter it is 80 periods ahead. The graph clearly shows that the optimal policy is steady through time and only approaching the terminal date the social planner chooses to stop curbing the level of emissions. This result holds for both the cases illustrated in the graphs and appears to be robust as we have obtained similar plots with several alternative configurations.[12]

4 Lagged State Observation

Another issue which is worth analyzing is what happens when the social planner needs choosing her optimal policy before observing the current emission and concentration levels. Indeed, data on concentration and emission levels are not immediately available, with a delay in their release which is particularly severe for the latter. Therefore, we assume that, as it takes time to collect such data and the social planner observes the state vector with a period lag, when choosing the effort level in t it knows the value of $\mathbf{z}_{t-1}$ but not that of $\mathbf{z}_t$.

Exploiting Theorem 4 and Lemma 7 in Vitale (2015) one can show that, in this case, for $\rho > 0$ the optimal mitigation policy in the steady state is found using in the optimal control rule presented in Proposition 1 in lieu of the state vector $\mathbf{z}_t$ a different vector $\breve{\mathbf{z}}_t$. Such vector corresponds to a transformation of the maximum likelihood (ML) estimate of $\mathbf{z}_t$, $\hat{\mathbf{z}}_t = \mathbf{A}\mathbf{z}_{t-1} + \mathbf{B}\mathbf{u}_{t-1}$, given by (see Appendix A.10)[13]

$$\breve{\mathbf{z}}_t \;=\; (\mathbf{I} \,-\, \rho \mathbf{N}\,\mathbf{\Pi})^{-1} \hat{\mathbf{z}}_t\,.$$

This implies that the optimal rule in matrix form in the lagged observation case is then equal to $u_t = \mathbf{K}\breve{\mathbf{z}}_t$. Notice that, given the expression for $\breve{\mathbf{z}}_t$, we can define the adjusted vector $\mathbf{K}_I = \mathbf{K}(\mathbf{I} \,-\, \rho\mathbf{N}\,\mathbf{\Pi})^{-1}$, so that the optimal policy can be written in terms of the ML estimate,

[12] These results are reported in the web Appendix.

[13] As mentioned in Section 2, for $\rho \downarrow 0$ the social planner solves a standard discounted linear quadratic gaussian (DLQG) problem. In this case the certainty equivalence principle can be applied and to find the optimal mitigation policy is sufficient to replace in the expression presented in Proposition 1 the state variables, p_t and e_t, with their ML estimates, $\hat{p}_t$ and $\hat{e}_t$. Thus, the optimal mitigation policy in matrix form is $u_t = \mathbf{K}\hat{\mathbf{z}}_t$, where $\mathbf{K} = (\kappa_p \;\; \kappa_e)$ and κ_p and κ_e are the coefficients derived in Proposition 1.

$u_t = \mathbf{K}_I \hat{\mathbf{z}}_t$.[14]

[Figures 5 and 6 about here.]

In Figures 5 and 6 we compare the steady state of the full observation case discussed in Section 3 with that of the lagged observation case. In Figure 5 we compare the coefficients κ_p and κ_e, and the optimal abatement effort u_t, evaluated in the case of full observation, with the adjusted coefficients κ_p^I and κ_e^I, and the optimal abatement effort u_t^I in the case of lagged observation. We see that the mitigation policy is more aggressive when the social planner observes the emission and concentration levels with a period lag (in absolute value κ_p^I, κ_e^I and u_t^I are larger than κ_p, κ_e and u_t for any positive value of ρ). As the state of the world becomes more uncertain the social planner chooses to curb more the emission level.

In Figure 6 we compare the unconditional standard deviation of the concentration and emission levels, $\sigma(p_t)$ and $\sigma(e_t)$, under full and lagged state observation. Despite a more aggressive policy the unconditional standard deviations of the concentration and emission levels are larger in the latter scenario, because the social planner faces a more uncertain environment.[15] However, as ρ increases and the social planner becomes more risk-averse she turns extremely aggressive in the lagged observation case, so that the difference between the two scenarios reduces. Finally, it is worth noticing that, given the estimated values for e_t and p_t in 2012, the ranges of values taken by $\sigma(p_t)$ and $\sigma(e_t)$ in Figure 6 are realistic.

5 A Robust mitigation Policy

The optimal control rule we have obtained is closely related to a robust decision rule which applies to a specific Linear Quadratic Gaussian (LQG) formulation. In particular, let us assume the state vector $\mathbf{z}_t$ respects the linear equation (2.4) and that the cost function is still given by (2.5). However, the social planner possesses quadratic preferences, so that in any period t

[14]Similar results would hold for a generalization of this formulation to imperfect state observation. In this case, the social planner observes in t a noisy signal of the state vector, $\mathbf{y}_t = \mathbf{H}\mathbf{z}_{t-1} + \boldsymbol{\zeta}_t$, with $\boldsymbol{\zeta}_t$ a white noise process independent of $\boldsymbol{\epsilon}_t$ ($\boldsymbol{\zeta}_t \sim N(\mathbf{0}, \mathbf{M})$).

[15]For the derivation under lagged state observation of the unconditional standard deviation in steady state of the concentration level, $\sigma(p_t)$, and the emission level, $\sigma(e_t)$, see Appendix A.11.

she chooses the optimal control solving the following program

$$\min_{\{u_t\}_{t=0}^{\infty}} \; E_t \left[\sum_{i=0}^{\infty} \delta^i c_{t+i} \right] , \tag{5.1}$$

$$\text{s.t.} \qquad \mathbf{z}_{t+1} = \mathbf{A}\,\mathbf{z}_t + \mathbf{B}\,u_t + \boldsymbol{\epsilon}_{t+1} .$$

This is a standard LQG problem. As this Section borrows heavily from [**?**] we employ their notation. Thus, we rewrite equation (2.4) as follows

$$\mathbf{z}_{t+1} \;=\; \mathbf{A}\,\mathbf{z}_t + \mathbf{B}\,u_t + \mathbf{C}\,\boldsymbol{\xi}_{t+1} , \quad \text{where } \boldsymbol{\xi}_{t+1} \sim N(\mathbf{0}, \mathbf{I}) \tag{5.2}$$

and $\mathbf{C}$ is the Cholesky decomposition of the matrix $\mathbf{N}$ (so that $\mathbf{N} = \mathbf{C}\mathbf{C}'$).

The social planner may suspect that equation (5.2) is not correct and it is just an approximation of the actual law of motion for the state vector (i.e. it represents an *approximating* model). In particular, the social planner may suspect that the law of motion corresponds to a *distorted* version of equation (5.2),

$$\mathbf{z}_{t+1} \;=\; \mathbf{A}\,\mathbf{z}_t + \mathbf{B}\,u_t + \mathbf{C}\,(\check{\boldsymbol{\xi}}_{t+1} + \mathbf{w}_{t+1}) , \tag{5.3}$$

where $\check{\boldsymbol{\xi}}_{t+1} \sim N(\mathbf{0}, \mathbf{I})$ and $\mathbf{w}_{t+1}$ is some unspecified process given by some measurable (non-necessarily linear) function of the state vector history (i.e. there exists $\mathbf{g}_t$ such that $\mathbf{w}_{t+1} = \mathbf{g}_t(\mathbf{z}_t, \mathbf{z}_{t-1}, \dots)$).

The social planner aims at choosing a mitigation policy which works for any alternative distorted model (5.3), as long as the *discrepancy* (i.e. the statistical or probabilistic distance) between the approximating and distorted models is not too large. To measure such discrepancy one relies on the concept of *conditional relative entropy.* In particular, let $f(\mathbf{z}_{t+1} \mid \mathbf{z}_t)$ denote the conditional transition density for the state vector according to the distorted model (5.3) and let $f_0(\mathbf{z}_{t+1} \mid \mathbf{z}_t)$ be the conditional transition density for the state vector according to the approximating model (5.2). One defines the conditional relative entropy as follows

$$I(f_0, f)(\mathbf{z}_t) \;\equiv\; \int \log\left(\frac{f(\mathbf{z}_{t+1} \mid \mathbf{z}_t)}{f_0(\mathbf{z}_{t+1} \mid \mathbf{z}_t)} \right) f(\mathbf{z}_{t+1} \mid \mathbf{z}_t)\, d\mathbf{z}_t .$$

As this is the expected *log-likelihood* ratio, this conditional relative entropy measures the probabilistic distance between the distorted and the approximating model. Under normality

Hansen and Sargent shows that

$$I(f_0, f)(\mathbf{z}_t) \;=\; \frac{1}{2}\,\mathbf{w}'_{t+1}\,\mathbf{w}_{t+1}\,.$$

As an *intertemporal* measure of distortion Hansen and Sargent employ the aggregate value

$$\mathcal{R}_t \;\equiv\; 2\,E_0\left[\sum_{i=0}^{\infty}\delta^i\, I(f_0, f)(\mathbf{z}_{t+1})\right] \;=\; E_0\left[\sum_{i=0}^{\infty}\delta^i\,\mathbf{w}'_{t+i}\,\mathbf{w}_{t+i}\right]\,.$$

Then, they consider all distorted models (5.3) alternative to the approximating model (5.2) for which $\mathcal{R}_t \leq \omega$, where ω is a maximum discrepancy value, representing an upper bound on the mis-specification of the approximating model.

In other words, following Hansen and Sargent, one can envision a situation in which the social planner assumes that data are generated by model (5.2) and suspects that they are actually generated by a distorted model (5.3) which is not too *far* from the approximating one. In measuring their distance she refers to the intertemporal conditional entropy $\mathcal{R}_t$. A *robust* control rule is then one which works for *all* distorted models for which $\mathcal{R}_t \leq \omega$. More precisely, the selection criterion proposed by Hansen and Sargent to define a robust control rule is particularly demanding, in that it requires that the social planner chooses the control rule which minimizes the expected aggregate cost of the *worst* distorted model (among all admissible ones). Formally, a robust control rule solves the following problem

$$\begin{aligned} &\min_{\{u_t\}_{t=0}^{\infty}}\;\max_{\{\mathbf{w}_{t+1}\}_{t=0}^{\infty}}\; E_t\left[\sum_{i=0}^{\infty}\delta^i c_{t+i}\right], \qquad (5.4)\\ \text{s.t.}\quad & \mathbf{z}_{t+1} \;=\; \mathbf{A}\,\mathbf{z}_t \;+\; \mathbf{B}\,u_t \;+\; \mathbf{C}\,(\check{\boldsymbol{\xi}}_{t+1} \;+\; \mathbf{w}_{t+1})\,,\\ & \mathcal{R}_t \;\leq\; \omega. \end{aligned}$$

This means that first among all alternative models the social planner chooses the worst-one, i.e. the one which maximizes her expected aggregate cost, and second she selects the optimal control rule which minimizes her aggregate cost within this worst-case scenario.

Importantly, it can be shown that the solution to problem (5.4) coincides to that of the recursive optimization (2.7) for a specific parametric choice (i.e. for a specific choice of ρ given ω). In other words, our risk-sensitive optimal control rule obtained from Proposition 1 is also a robust optimal control rule *à la* Hansen and Sargent. The correspondence between the two formulations extends further. In fact, Hansen and Sargent indicate that their robust optimal control problem admits a solution insofar $\omega \leq \bar{\omega}$, where $\bar{\omega}$ is a maximum possible level for the degree of uncertainty of the social planner on the model mis-specification. This condition is analogous to the requirement that the risk-enhancement coefficient ρ in the recursive optimization (2.7) is not too large, so that the second order condition that $(\delta\mathbf{\Pi})^{-1} - \rho\mathbf{N}$ is positive definite is satisfied.

This suggests that our analysis can have a double interpretation. It can be considered an investigation of the impact on the optimal mitigation policy of either the risk-aversion of the social planner or her uncertainty on the model governing the dynamics of the concentration and emission of GHGs. In this respect, our analysis could be considered complementary to the contribution of [**?**], who exploit Hansen and Sargent's robust optimal control methodology to recommend the precautionary principle in the conduct of climate change policy. Our analysis, however, differs from theirs in several dimensions.

Indeed, differently from [**?**] we consider a formulation in which the emission and concentration levels are both subject to stochastic shocks so that in our model the social planner does not have full control on emission flows. Secondly, in our cost function we model the direct cost of abatement while [**?**] accommodate the indirect cost of abatement associated with the reduction in output alongside the cost of investing in a damage control technology. Thirdly, we consider a discrete-time model which allows for a terminal date for the intervention horizon of the social planner. Fourthly, the stylized model we consider is based on the presumption that a mitigation effort needn't impair economic activity. Finally, and most importantly, when we interpret our model in terms of risk-sensitive preferences, we are also able to consider an extension in which the social planner only observes the state variables with a time lag, a scenario that [**?**] do not investigate since their methodology does not allow for learning on the mis-specification of the approximating model.[16] This differentiation brings about some differences in results. Notably, while we show that larger uncertainty has always a positive impact on the optimal mitigation effort, in their formulation this result is valid only when the optimal level of investment is sufficiently low. Nonetheless the effect of risk-aversion on the optimal

[16]In principle, by observing the history of the state vector, $\{\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots, \mathbf{z}_0\}$, it should be possible to back out the sequence of error terms in the approximating model, $\{\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}, \cdots, \boldsymbol{\xi}_0\}$, and consequently make some inference on their probabilistic properties. This should permit the social planner to learn about the mis-specification of the approximating model (5.2) and reduce over time the degree of uncertainty she faces. Unfortunately, Hansen and Sargent's robustness methodology does permit to do that. On the contrary, such a limitation is absent in the DLEQG formulation, which also allows for the possibility that the social planner observes imperfectly the state vector.

mitigation effort is similar to that of more uncertainty on the impact of the investment in the damage-control technology that [?] investigate.

6 Concluding Remarks

Our paper contributes to the current debate on how risk and uncertainty should be taken into account when investigating optimal climate change policy. We characterize the optimal mitigation policy of a social planner who assigns greater importance to extreme but rare catastrophic events and minimizes her cost function when the concentration of GHGs in the atmosphere is equal to zero. We find that risk-sensitive preferences can significantly increase the mitigation effort so that the social planner's best reaction to the worst possible shocks to the environment consists in a very early and aggressive mitigation policy. Moreover, the social planner's preferences and the strong inertia of GHGs in the atmosphere entail an *inverted policy ramp* in that the mitigation effort starts very aggressively and then it slows down gradually. We also observe that in the first phase of the convergence process such dynamics brings about a significant level of carbon sequestration.

Our results seem to provide an affirmative answer to the increasing requests for earlier and more energetic interventions in the fight against climate change. Indeed, the precautionary principle, advocated by others via either a small discount rate [?] or fat-tailed probability distributions [?] or uncertainty over the dynamics of the environmental conditions [?], is imposed in this paper by a pessimistic optimal choice mechanism.

References

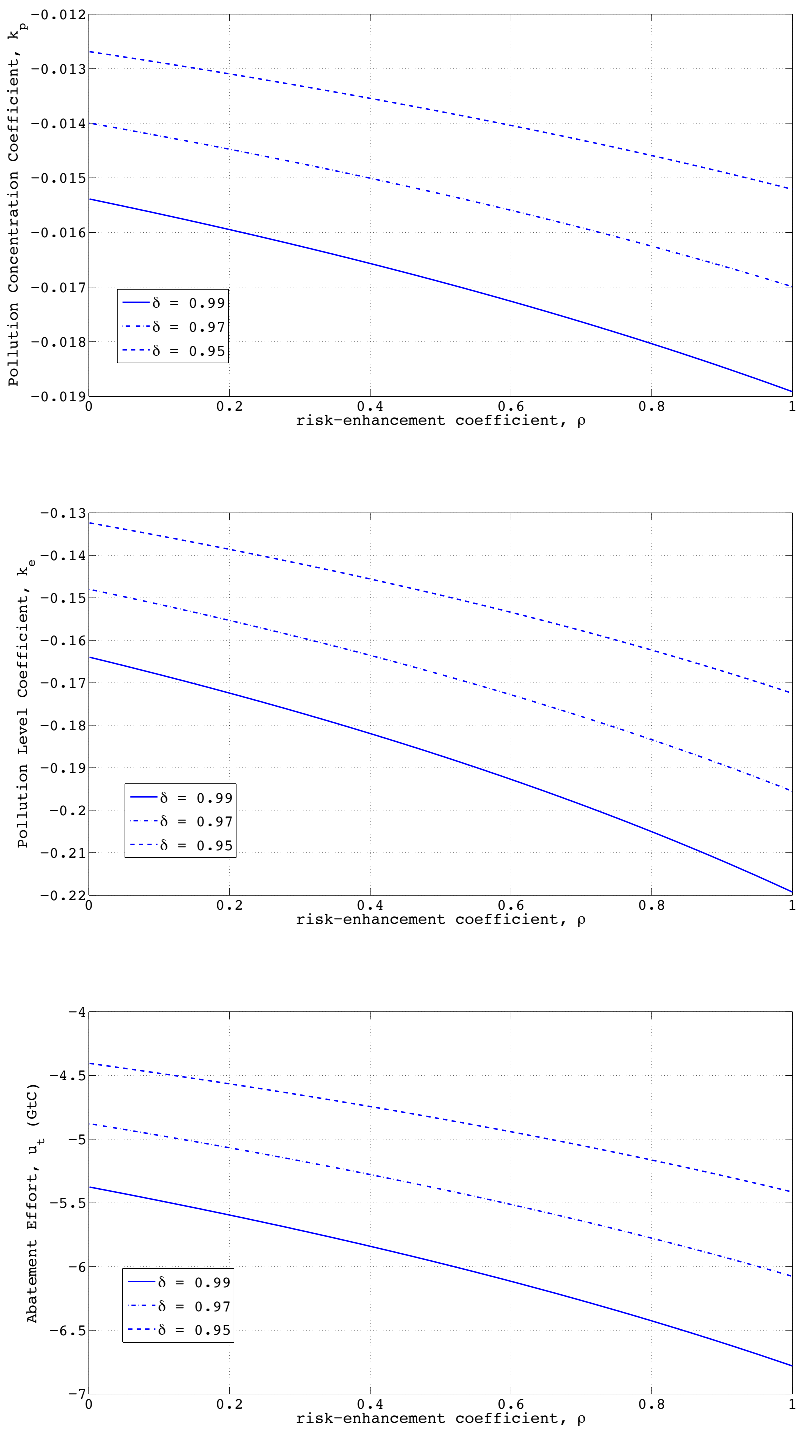

Figure 1: Dependence of κ_p (top panel), κ_e (middle panel), and the optimal abatement effort, u_t (bottom panel), given the concentration and emission levels in 2012 ($e_0 = 9.7$ and $p_0 = 246$), on ρ, the risk-enhancement coefficient, for $\delta = 0.95$, $\delta = 0.97$ and $\delta = 0.99$, when $\alpha = 30$, $\beta = 0.01115$, $\gamma = 0.9917$, $\sigma_e^2 = 0.01$, $\sigma_p^2 = 0.0549$. In the top and middle panels κ_p and κ_e are the coefficients on the concentration level and on the emission level according to the optimal mitigation policy in Proposition 1.

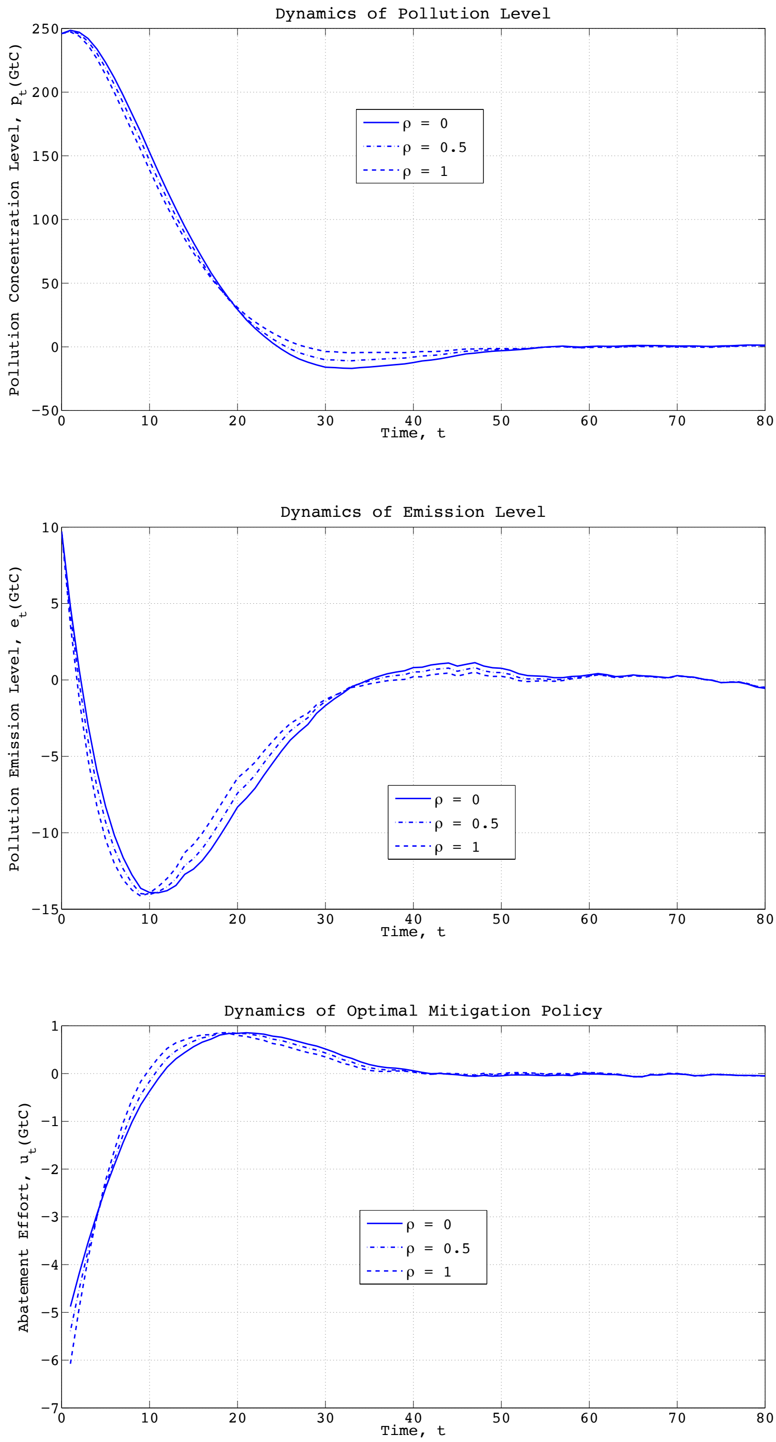

Figure 2: The simulation of the dynamics of the concentration level p_t (top panel), the emission level e_t (middle panel), and the optimal abatement effort u_t (bottom panel), over 80 periods, given the initial concentration and emission levels in 2012 ($p_0 = 246$ and $e_0 = 9.7$), for $\rho = 0$, $\rho = 0.5$ and $\rho = 1$, when $\delta = 0.97$, $\alpha = 30$, $\beta = 0.01115$, $\gamma = 0.9917$, $\sigma_e^2 = 0.01$, $\sigma_p^2 = 0.0549$.

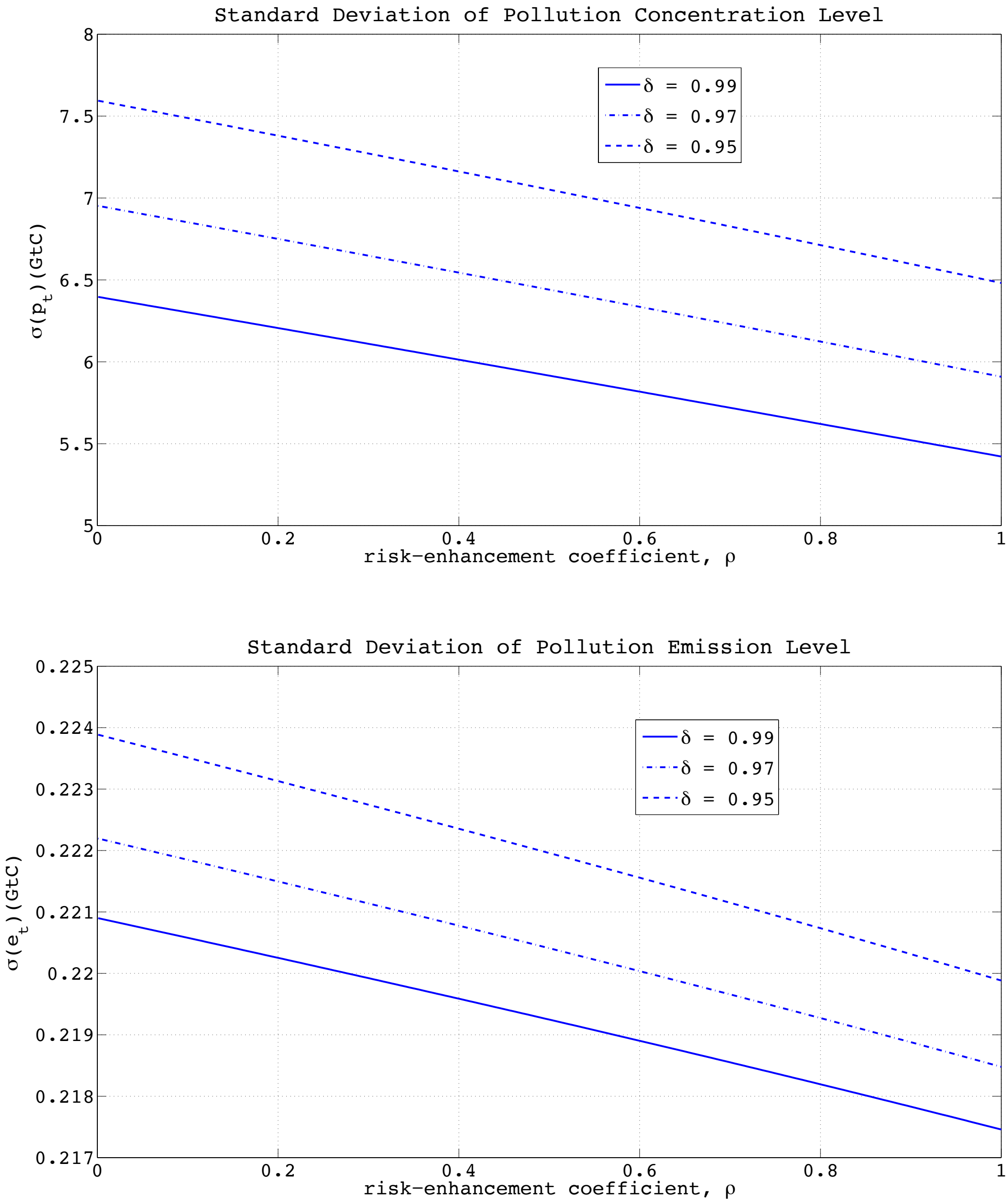

Figure 3: Dependence of the unconditional standard deviation of the concentration (top panel) and emission levels (bottom panel) on ρ, the risk-enhancement coefficient, for $\delta = 0.99$, $\delta = 0.97$ and $\delta = 0.95$, when $\alpha = 30$, $\beta = 0.01115$, $\gamma = 0.9917$, $\sigma_e^2 = 0.01$, $\sigma_p^2 = 0.0549$.

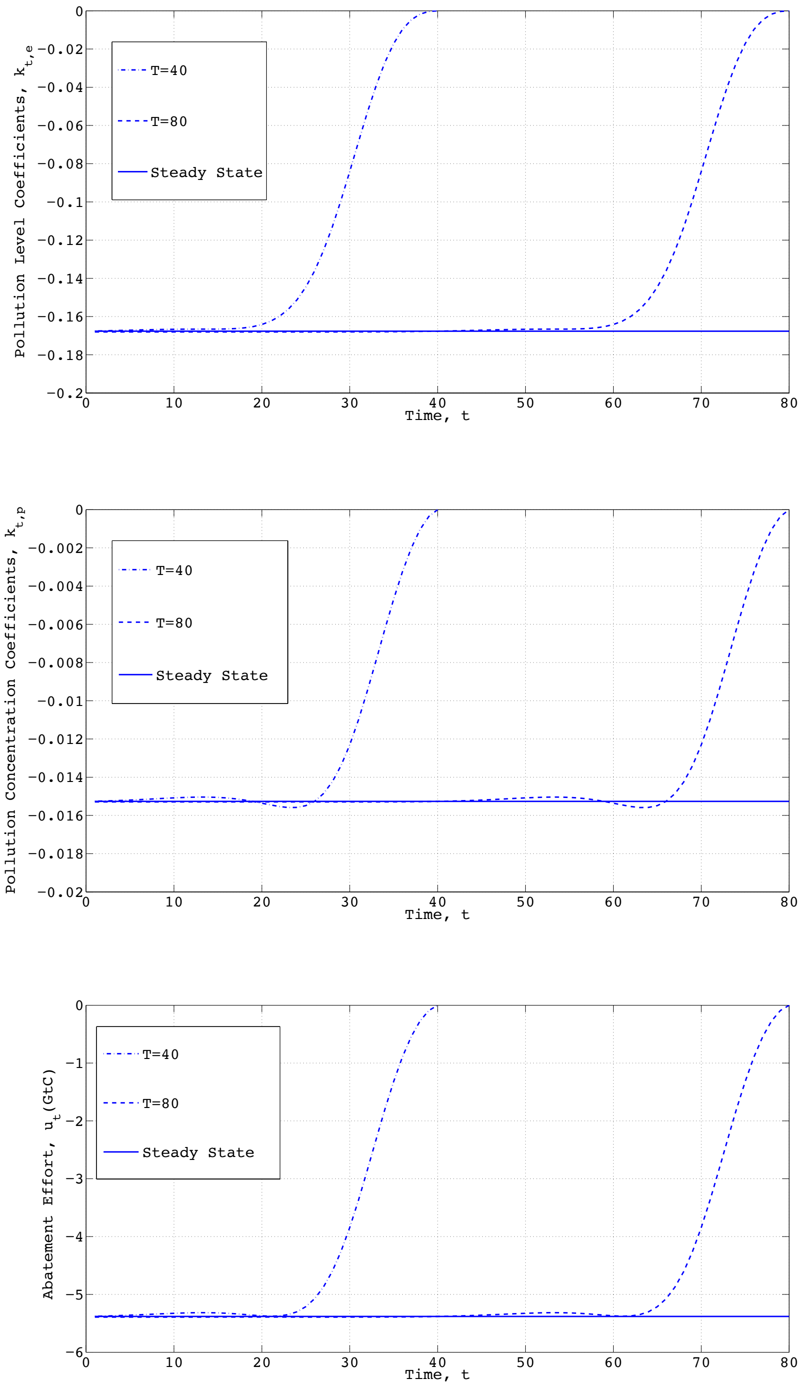

Figure 4: Dynamics of κ_e (top panel), κ_p (middle panel), and the optimal abatement effort (bottom panel) for two values of the policy implementation period, $T = 40$ and $T = 80$, when $\alpha = 30$, $\beta = 0.01115$, $\delta = 0.97$, $\gamma = 0.9917$, $\rho = 0.5$, $\sigma_e^2 = 0.0549$, $\sigma_p^2 = 0.01$. In the top and middle panels κ_p and κ_e are the coefficients on the concentration level and on the emission level according to the optimal mitigation policy in Proposition 1. In the bottom panel the optimal mitigation effort is calculated assuming that the pollution, p_t, and emission levels, e_t, are equal to those observed in 2012, that is equal respectively to 246 and 9.7.

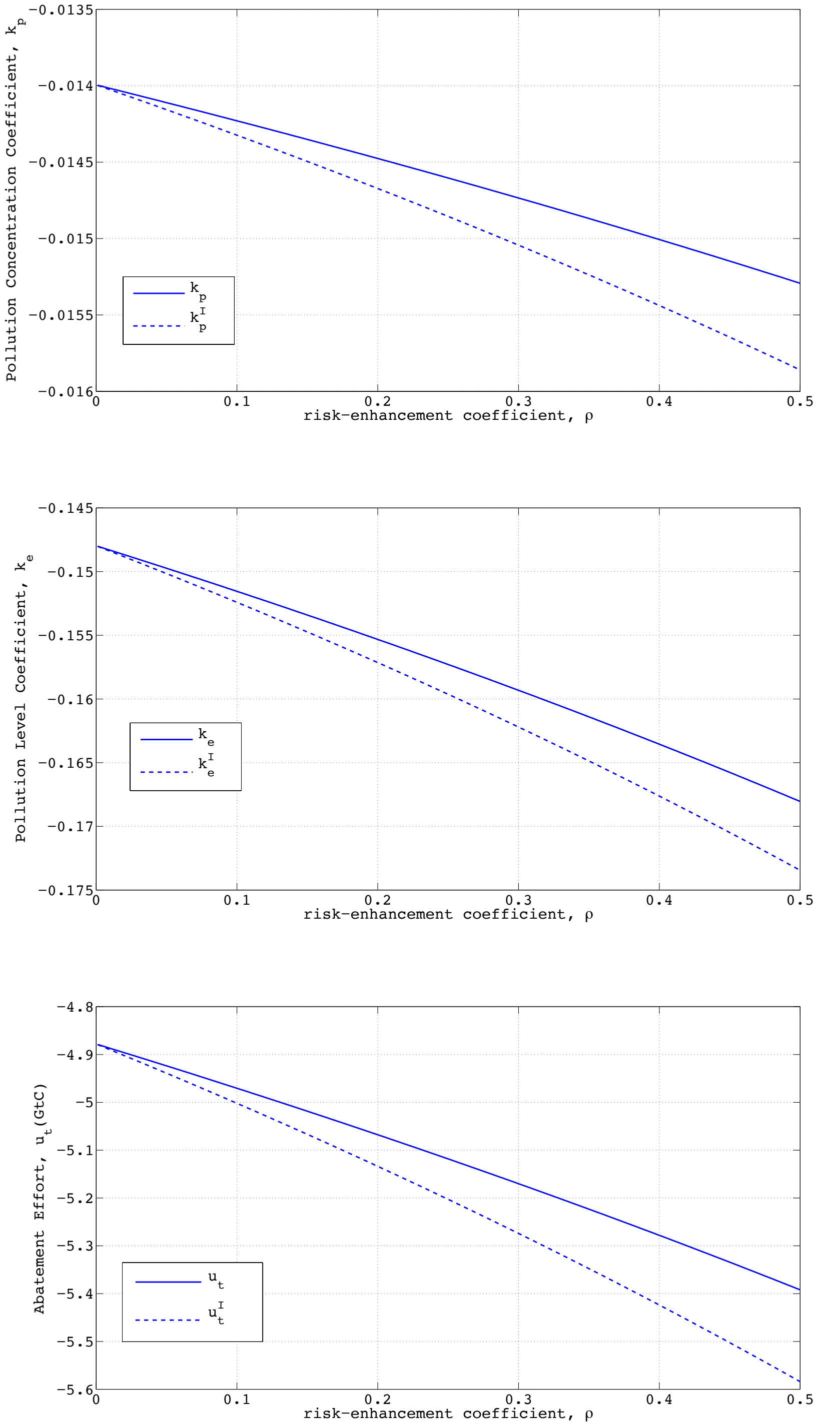

Figure 5: Dependence of κ_p (top panel), κ_e (middle panel), and the optimal abatement effort, u_t (bottom panel), given the concentration and emission levels in 2012 ($e_0 = 9.7$ and $p_0 = 246$), on ρ, the risk-enhancement coefficient, in the perfect and lagged observation scenario when $\alpha = 30$, $\beta = 0.01115$, $\delta = 0.97$, $\gamma = 0.9917$, $\sigma_e^2 = 0.01$ and $\sigma_p^2 = 0.0549$. In the top and middle panels κ_p and κ_e are the coefficients on the concentration level and on the emission level according to the optimal mitigation policy in Proposition 1.

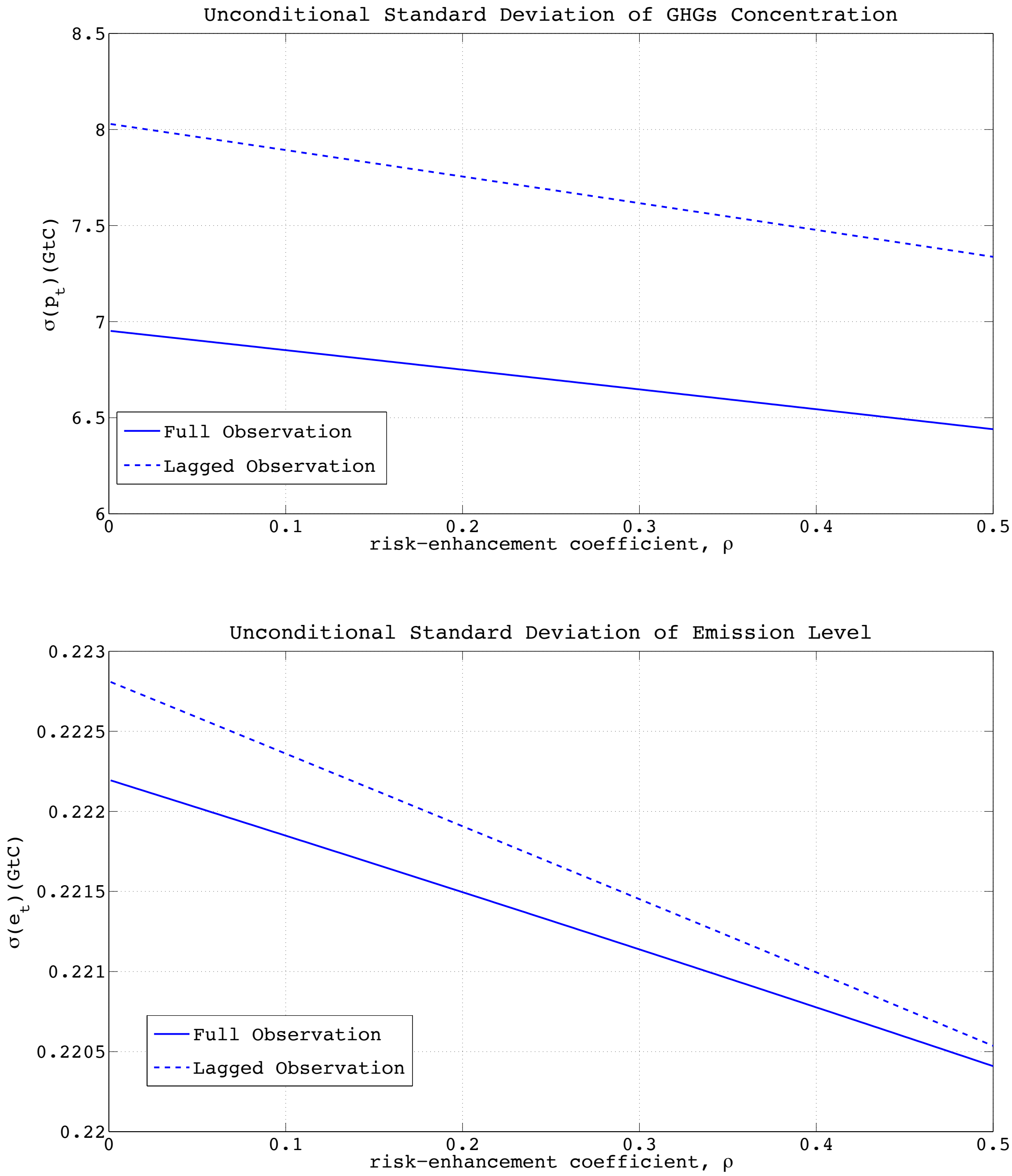

Figure 6: Dependence of the unconditional standard deviation of the concentration (top panel) and emission levels (bottom panel) on ρ, the risk-enhancement coefficient, in the perfect and lagged observation scenario when $\alpha = 30$, $\beta = 0.01115$, $\delta = 0.97$, $\gamma = 0.9917$, $\sigma_e^2 = 0.01$ and $\sigma_p^2 = 0.0549$.

A. Detailed Appendix

A.1. The Solution of the Recursive Optimization. Consider that we solve the recursion

$$\exp\left(\frac{\rho}{2}\boldsymbol{\mathcal{V}}_t\right) \quad = \quad \min_{u_t} E_t\left[\exp\left(\frac{\rho}{2}(c_t \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right], \tag{A.1}$$

where $c_t \ = \ \beta\, p_t^2 \ + \ \alpha\, u_t^2 \ + \ \eta_t$. Given that η_t is independent of $\boldsymbol{\epsilon}_t$, we find that

$$\exp\left(\frac{\rho}{2}\boldsymbol{\mathcal{V}}_t\right) \quad = \quad E_t\left[\exp\left(\frac{\rho}{2}\eta_t\right)\right] \times \min_{u_t} E_t\left[\exp\left(\frac{\rho}{2}(c_{o,t} \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right],$$

where $c_{o,t} = \beta\, p_t^2 \ + \ \alpha\, u_t^2$. Because η_t is normally distributed, we can write

$$\exp\left(\frac{\rho}{2}\boldsymbol{\mathcal{V}}_t\right) \quad = \quad \exp\left(\frac{\rho}{2}\frac{1}{4}\rho\sigma_\eta^2\right) \times \min_{u_t} E_t\left[\exp\left(\frac{\rho}{2}(c_{o,t} \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right]. \tag{A.2}$$

This means that the idiosyncratic shock, η_t, does not hinge on the optimal control rule. This is found by minimizing with respect to u_t the expected value of the exponential of $\frac{\rho}{2}(c_{o,t} + \delta\,\boldsymbol{\mathcal{V}}_{t+1})$. Because $c_{o,t}$ is a quadratic form in u_t and $\mathbf{z}_t$, we can rely on results in [**?**] pertaining to the analysis of discounted linear exponential quadratic Gaussian (DLEQG) problems *à la* [**?**].

A.2. Limit Properties of the Optimization Criterion. Consider that

$$\begin{aligned}
\exp\left(\frac{\rho}{2}\boldsymbol{\mathcal{V}}_t\right) &= \exp\left(\frac{\rho}{2}\frac{1}{4}\rho\sigma_\eta^2\right) \times \min_{u_t} E_t\left[\exp\left(\frac{\rho}{2}(c_{o,t} \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right] \iff \\
\frac{\rho}{2}\boldsymbol{\mathcal{V}}_t &= \frac{\rho}{2}\frac{1}{4}\rho\sigma_\eta^2 \ + \ \ln\left(\min_{u_t} E_t\left[\exp\left(\frac{\rho}{2}(c_{o,t} \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right]\right) \iff \\
\frac{\rho}{2}\boldsymbol{\mathcal{V}}_t &= \frac{\rho}{2}\frac{1}{4}\rho\sigma_\eta^2 \ + \ \min_{u_t}\ln\left(E_t\left[\exp\left(\frac{\rho}{2}(c_{o,t} \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right]\right) \iff \\
\frac{\rho}{2}\boldsymbol{\mathcal{V}}_t &= \frac{\rho}{2}\frac{1}{4}\rho\sigma_\eta^2 \ + \ \min_{u_t}\ln\left(\exp\left(\frac{\rho}{2}c_{o,t}\right) E_t\left[\exp\left(\delta\,\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right]\right) \iff \\
\frac{\rho}{2}\boldsymbol{\mathcal{V}}_t &= \frac{\rho}{2}\frac{1}{4}\rho\sigma_\eta^2 \ + \ \min_{u_t}\left(\frac{\rho}{2}c_{o,t} \ + \ \ln E_t\left[\exp\left(\delta\,\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right]\right) \iff
\end{aligned}$$

For $\rho > 0$ we have that

$$\frac{\rho}{2}\boldsymbol{\mathcal{V}}_t \ = \ \rho\,\min_{u_t}\left\{\frac{1}{2}\frac{1}{4}\rho\sigma_\eta^2 \ + \ \frac{1}{2}c_{o,t} + \frac{1}{\rho}\ln\left(E_t\left[\exp\left(\delta\,\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right]\right)\right\}$$

and hence that

$$\boldsymbol{\mathcal{V}}_t \quad = \quad \min_{u_t}\left\{\frac{1}{4}\rho\sigma_\eta^2 \ + \ c_{o,t} \ + \ 2\,\frac{1}{\rho}\ln\left(E_t\left[\exp\left(\delta\,\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right]\right)\right\}.$$

Consider that if $\boldsymbol{\mathcal{V}}_{t+1}$ is independent of ρ,

$$\lim_{\rho\downarrow 0} \frac{1}{\rho} \ln\left(E_t\left[\exp\left(\delta\,\frac{\rho}{2}\,\boldsymbol{\mathcal{V}}_{t+1}\right)\right]\right) \;=\; \lim_{\rho\downarrow 0} \delta\,\frac{1}{2}\,\frac{E_t\left[\exp\left(\delta\,\frac{\rho}{2}\,\boldsymbol{\mathcal{V}}_{t+1}\right)\cdot\boldsymbol{\mathcal{V}}_{t+1}\right]}{E_t\left[\exp\left(\delta\,\frac{\rho}{2}\,\boldsymbol{\mathcal{V}}_{t+1}\right)\right]} \;=\; \frac{1}{2}\,\delta\,E_t\left[\boldsymbol{\mathcal{V}}_{t+1}\right]\,,$$

where we have used the Hôpital's rule and moved the derivative operator inside the expectation operator. Noting that $\lim_{\rho\downarrow 0}\,\frac{1}{4}\,\rho\sigma_\eta^2 = 0$, this implies that

$$\lim_{\rho\downarrow 0} \boldsymbol{\mathcal{V}}_t \;=\; \min_{u_t}\left\{c_{o,t} \;+\; \delta\,E_t\left[\boldsymbol{\mathcal{V}}_{t+1}\right]\right\}\,,$$

with $\boldsymbol{\mathcal{V}}_t$ independent of ρ. If we have a terminal date in T, by definition $\boldsymbol{\mathcal{V}}_{T+1} = 0$ (i.e. independent of ρ). Then, by backward induction for $\rho\downarrow 0$ our recursive optimization converges to the Bellman equation of the corresponding DLQG problem.

A.3. The Recursive Optimization and Epstein-Zin Preferences. Suppose $\boldsymbol{\mathcal{U}}_t$ solves Epstein and Zin's recursion

$$\boldsymbol{\mathcal{U}}_t \;=\; \max\left\{(1-\delta)C_t^{1-\frac{1}{\theta}} \;+\; \delta\,E_t\,\left[\boldsymbol{\mathcal{U}}_{t+1}^{1-\chi}\right]^{\frac{1-\frac{1}{\theta}}{1-\chi}}\right\}^{\frac{1}{1-\frac{1}{\theta}}}\,,$$

where θ is the elasticity of inter-temporal substitution. Let $\theta = 1$. [**?**] shows that

$$\boldsymbol{\mathcal{U}}_t \;=\; \max\left\{C_t^{1-\delta}\,\left(E_t\,\left[\boldsymbol{\mathcal{U}}_{t+1}^{1-\chi}\right]\right)^{\left(\frac{\delta}{1-\chi}\right)}\right\}\,.$$

Taking logs,

$$\log\boldsymbol{\mathcal{U}}_t \;=\; \max\left\{(1-\delta)\,\log C_t \;+\; \frac{\delta}{1-\chi}\,\log E_t\,\left[\boldsymbol{\mathcal{U}}_{t+1}^{1-\chi}\right]\right\}\,,$$

or equivalently

$$\frac{\log\boldsymbol{\mathcal{U}}_t}{1-\delta} \;=\; \max\left\{\log C_t \;+\; \frac{\delta}{(1-\delta)(1-\chi)}\,\log E_t\,\left[\boldsymbol{\mathcal{U}}_{t+1}^{1-\chi}\right]\right\}\,.$$

We can re-write this as

$$-\frac{\log\boldsymbol{\mathcal{U}}_t}{1-\delta} \;=\; \min\left\{-\,\log C_t \;-\; \frac{\delta}{(1-\delta)(1-\chi)}\,\log E_t\,\left[\boldsymbol{\mathcal{U}}_{t+1}^{1-\chi}\right]\right\}\,.$$

For $\boldsymbol{\mathcal{V}}_t = -\frac{\log\boldsymbol{\mathcal{U}}_t}{1-\delta}$, we have that $-(1-\delta)\boldsymbol{\mathcal{V}}_t = \log\boldsymbol{\mathcal{U}}_t$, so that $\boldsymbol{\mathcal{U}}_{t+1} = \exp(-(1-\delta)\boldsymbol{\mathcal{V}}_{t+1})$ and

$$\boldsymbol{\mathcal{U}}_{t+1}^{1-\chi} \;=\; \left(\exp(-(1-\delta)\boldsymbol{\mathcal{V}}_{t+1})\right)^{1-\chi} \;=\; \exp(-(1-\delta)(1-\chi)\boldsymbol{\mathcal{V}}_{t+1})\,.$$

Setting $\rho' = -2(1-\delta)(1-\chi)$, we can write

$$\boldsymbol{\mathcal{V}}_t \;=\; \min\left\{-\log C_t \;+\; \delta\,\frac{2}{\rho'}\,\log E_t\,\left[\exp\left(\frac{\rho'}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right]\right\}\,,$$

Define $\rho = \rho'/\delta = \ 2\left(\frac{1}{\delta} - 1\right)(\chi - 1)$ and notice that

$$\begin{aligned} \boldsymbol{\mathcal{V}}_t &= \min\left\{ -\log C_t \ + \ \frac{2}{\rho}\ \log E_t\ \left[\exp\left(\delta\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right] \right\} \quad \Longleftrightarrow \\ \frac{\rho}{2}\,\boldsymbol{\mathcal{V}}_t &= \min\left\{ -\frac{\rho}{2}\ \log C_t \ + \ \log E_t\ \left[\exp\left(\delta\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right] \right\} . \end{aligned}$$

Suppose that $-\log C_t$ equal to a quadratic form in the control and state vectors, u_t and $\mathbf{z}_t$, c_t, then

$$\begin{aligned} \exp\left(\frac{\rho}{2}\,\boldsymbol{\mathcal{V}}_t\right) &= \exp\left(\min\left\{ \frac{\rho}{2}\,c_t \ + \ \log E_t\ \left[\exp\left(\delta\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right] \right\}\right) \\ &= \min\left\{ \exp\left(\frac{\rho}{2}\,c_t \ + \ \log E_t\ \left[\exp\left(\delta\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right]\right)\right\} \\ &= \min\left\{ \exp\left(\frac{\rho}{2}\,c_t\right)\ \exp\left(\log E_t\ \left[\exp\left(\delta\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right]\right)\right\} \\ &= \min\left\{ \exp\left(\frac{\rho}{2}\,c_t\right)\ E_t\ \left[\exp\left(\delta\frac{\rho}{2}\boldsymbol{\mathcal{V}}_{t+1}\right)\right]\right\} \\ &= \min\left\{ E_t\left[\exp\left(\frac{\rho}{2}\,(c_t \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right]\right\} , \end{aligned}$$

which corresponds to the recursive optimization we employ if $\sigma_\eta^2 = 0$.

A.4. The Derivation of Lemma 1. Consider that

$$\begin{aligned} E_t\left[\exp\left(\frac{\rho}{2}(c_{o,t} \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right] &= (2\pi)^{-1}\det(\mathbf{N})^{-1/2}\int \exp\left(\frac{\rho}{2}(c_{o,t} \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1}) \ - \ \frac{1}{2}\,\boldsymbol{\epsilon}'_{t+1}\mathbf{N}^{-1}\boldsymbol{\epsilon}_{t+1}\right)\,d\boldsymbol{\epsilon}_{t+1} \\ &= (2\pi)^{-1}\det(\mathbf{N})^{-1/2}\int \exp\left(\rho\,\frac{\boldsymbol{\mathcal{S}}_t}{2}\right)\,d\boldsymbol{\epsilon}_{t+1}\,, \end{aligned}$$

where $\boldsymbol{\mathcal{S}}_t = c_{o,t} - \frac{1}{\rho}\boldsymbol{\epsilon}'_{t+1}\mathbf{N}^{-1}\boldsymbol{\epsilon}_{t+1} + \delta\boldsymbol{\mathcal{V}}_{t+1}$. Then,

$$\min_{u_t}\ E_t\left[\exp\left(\frac{\rho}{2}(c_{o,t} \ + \ \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right] \quad = \quad (2\pi)^{-1}\det(\mathbf{N})^{-1/2}\min_{u_t}\int \exp\left(\rho\,\frac{\boldsymbol{\mathcal{S}}_t}{2}\right)\,d\boldsymbol{\epsilon}_{t+1}\,.$$

If $\boldsymbol{\mathcal{V}}_{t+1} = \lambda_{t+1} + \mathbf{z}'_{t+1}\mathbf{\Pi}_{t+1}\mathbf{z}_{t+1}$, the function $-\rho\boldsymbol{\mathcal{S}}_t$ is a quadratic form in u_t and $\boldsymbol{\epsilon}_{t+1}$, which we can write as $\mathbf{S}_{uu}u_t^2 + 2u_t\mathbf{S}_{u\boldsymbol{\epsilon}}\boldsymbol{\epsilon}_{t+1} + \boldsymbol{\epsilon}'_{t+1}\mathbf{S}_{\boldsymbol{\epsilon\epsilon}}\boldsymbol{\epsilon}_{t+1}$, with $\mathbf{S}_{\boldsymbol{\epsilon\epsilon}} = \mathbf{N}^{-1} - \delta\rho\mathbf{\Pi}_{t+1}$. We can apply Lemma 3 in [**?**]. From its proof we know that

$$\min_{u_t}\int \exp\left(\rho\,\frac{\boldsymbol{\mathcal{S}}_t}{2}\right)\,d\boldsymbol{\epsilon}_{t+1} \quad = \quad 2\pi\det(\mathbf{N}^{-1} - \delta\rho\mathbf{\Pi}_{t+1})^{-1/2}\times\exp\left(\frac{\rho}{2}\,\min_{u_t}\,\max_{\boldsymbol{\epsilon}_{t+1}}\,\boldsymbol{\mathcal{S}}_t\right).$$

Notice that $\mathbf{N}^{-1}-\delta\rho\mathbf{\Pi}_{t+1} = \mathbf{N}^{-1}(\mathbf{I}-\delta\rho\mathbf{N}\mathbf{\Pi}_{t+1})$, so that $\det(\mathbf{N}^{-1}-\delta\rho\mathbf{\Pi}_{t+1}) = \det(\mathbf{I}-\delta\rho\mathbf{N}\mathbf{\Pi}_{t+1})/\det(\mathbf{N})$.

Therefore,

$$E_t\left[\exp\left(\frac{\rho}{2}(c_{o,t} + \delta\,\boldsymbol{\mathcal{V}}_{t+1})\right)\right] = \det(\mathbf{I} - \delta\,\rho\mathbf{N}\,\mathbf{\Pi}_{t+1})^{-1/2}\min_{u_t}\int\exp\left(\rho\,\frac{\boldsymbol{\mathcal{S}}_t}{2}\right)d\boldsymbol{\epsilon}_{t+1}$$
$$= \exp\left(\frac{\rho}{2}\left[\nu_t + \min_{u_t}\max_{\boldsymbol{\epsilon}_{t+1}}\boldsymbol{\mathcal{S}}_t\right]\right),$$

where $\nu_t = \frac{2}{\rho}\log(\det(\mathbf{I} - \delta\,\rho\mathbf{N}\,\mathbf{\Pi}_{t+1})^{-1/2}) = -\frac{1}{\rho}\log(\det(\mathbf{I} - \delta\,\rho\mathbf{N}\,\mathbf{\Pi}_{t+1}))$.

In brief, from equation (A.2) we conclude that $\boldsymbol{\mathcal{V}}_t = \frac{1}{4}\rho\sigma_\eta^2 + \nu_t + \min_{u_t}\max_{\boldsymbol{\epsilon}_{t+1}}\boldsymbol{\mathcal{S}}_t$.

A.5. The Derivation of Proposition 1. For $\boldsymbol{\mathcal{V}}_{t+1} = \lambda_{t+1} + \mathbf{z}'_{t+1}\mathbf{\Pi}_{t+1}\mathbf{z}_{t+1}$ it follows that

$$\min_{u_t}\max_{\boldsymbol{\epsilon}_{t+1}}\boldsymbol{\mathcal{S}}_t = \min_{u_t}\left\{\max_{\boldsymbol{\epsilon}_{t+1}}\left[c_{o,t} - \frac{1}{\rho}\,d_{t+1} + \delta\lambda_{t+1} + \delta\,\mathbf{z}'_{t+1}\mathbf{\Pi}_{t+1}\mathbf{z}_{t+1}\right]\right\}$$
$$= \delta\lambda_{t+1} + \min_{u_t}\left\{\max_{\boldsymbol{\epsilon}_{t+1}}\left[c_{o,t} - \frac{1}{\rho}\,d_{t+1} + \delta\,\mathbf{z}'_{t+1}\mathbf{\Pi}_{t+1}\mathbf{z}_{t+1}\right]\right\}$$
$$= \delta\lambda_{t+1} + \mathbf{z}'_t\mathbf{\Pi}_t\mathbf{z}_t\,.$$

Hence, $\boldsymbol{\mathcal{V}}_t = \lambda_t + \mathbf{z}'_t\mathbf{\Pi}_t\mathbf{z}_t$, with $\lambda_t = \frac{1}{4}\rho\sigma_\eta^2 + \nu_t + \delta\lambda_{t+1}$.Given that $c_{o,t} = \mathbf{z}_t\mathbf{R}\mathbf{z}_t + \mathbf{Q}u_t^2$ and that $\mathbf{z}_{t+1} = \mathbf{A}\mathbf{z}_t + \mathbf{B}u_t + \boldsymbol{\epsilon}_{t+1}$, using Lemma 4 and Theorem 2 in [**?**], one finds that

$$\mathbf{\Pi}_t = \mathbf{R} + \mathbf{A}'\widetilde{\mathbf{\Pi}}_{t+1}\mathbf{A} - \mathbf{A}'\widetilde{\mathbf{\Pi}}_{t+1}\mathbf{B}(\mathbf{Q} + \mathbf{B}'\widetilde{\mathbf{\Pi}}_{t+1}\mathbf{B})^{-1}\mathbf{B}'\widetilde{\mathbf{\Pi}}_{t+1}\mathbf{A}\,, \tag{A.3}$$

$$\text{with}\quad \widetilde{\mathbf{\Pi}}_{t+1} = ((\delta\mathbf{\Pi}_{t+1})^{-1} - \rho\,\mathbf{N})^{-1}\,. \tag{A.4}$$

In addition, from the same results one immediately see that the saddle point for the discounted total stress is found for $u_t = \mathbf{K}_t\mathbf{z}_t$ with

$$\mathbf{K}_t = (\mathbf{Q} + \mathbf{B}'\widetilde{\mathbf{\Pi}}_{t+1}\mathbf{B})^{-1}\mathbf{B}'\widetilde{\mathbf{\Pi}}_{t+1}\mathbf{A}\,. \tag{A.5}$$

In steady state, $\widetilde{\mathbf{\Pi}}_{t+1} = \widetilde{\mathbf{\Pi}}$ and $\mathbf{\Pi}_{t+1} = \mathbf{\Pi}_t = \mathbf{\Pi}$, so that equations (A.3),(A.4) and (A.5) corresponds to the expressions in Proposition 1. In addition, because $\lambda_t = \nu_t + \delta\lambda_{t+1}$, $\boldsymbol{\mathcal{V}}_t = \mathbf{z}'_t\mathbf{\Pi}\mathbf{z}_t + \lambda$, where

$$\lambda = \frac{1}{1-\delta}\left(\frac{1}{4}\rho\sigma_\eta^2 - \frac{1}{\rho}\log(\det(\mathbf{I} - \delta\,\rho\mathbf{N}\,\mathbf{\Pi}))\right).$$

Then, suppose $\widetilde{\mathbf{\Pi}} = \begin{pmatrix} \tilde{\pi}_1 & \tilde{\pi}_{1,2} \\ \tilde{\pi}_{1,2} & \tilde{\pi}_2 \end{pmatrix}$. Given $\mathbf{A}$ and $\mathbf{B}$,

$$\begin{aligned} \mathbf{A}'\,\widetilde{\mathbf{\Pi}}\mathbf{B} &= \begin{pmatrix} \gamma(\tilde{\pi}_1 + \tilde{\pi}_{1,2}) \\ \tilde{\pi}_1 + 2\tilde{\pi}_{1,2} + \tilde{\pi}_2 \end{pmatrix}, \\ \mathbf{B}'\,\widetilde{\mathbf{\Pi}}\mathbf{B} &= \tilde{\pi}_1 + 2\tilde{\pi}_{1,2} + \tilde{\pi}_2 \,. \end{aligned}$$

Then, given $\mathbf{Q}$, exploiting Theorem 2 in [**?**], we find that

$$\begin{aligned} \mathbf{K}' &= -\,\mathbf{A}'\,\widetilde{\mathbf{\Pi}}\mathbf{B}\,(\mathbf{Q} + \mathbf{B}'\widetilde{\mathbf{\Pi}}\mathbf{B})^{-1} \\ &= -\,\frac{1}{\alpha + \tilde{\pi}_1 + 2\,\tilde{\pi}_{1,2} + \tilde{\pi}_2} \begin{pmatrix} \gamma(\tilde{\pi}_1 + \tilde{\pi}_{1,2}) \\ \tilde{\pi}_1 + 2\tilde{\pi}_{1,2} + \tilde{\pi}_2 \end{pmatrix}. \end{aligned}$$

This corresponds to $\mathbf{K} = (\kappa_p\ \kappa_e)$, with

$$\begin{aligned} \kappa_p &= -\,\gamma\left(1 - \frac{\alpha + \tilde{\pi}_{1,2} + \tilde{\pi}_2}{\alpha + \tilde{\pi}_1 + 2\,\tilde{\pi}_{1,2} + \tilde{\pi}_2}\right), \\ \kappa_e &= -\left(1 - \frac{\alpha}{\alpha + \tilde{\pi}_1 + 2\,\tilde{\pi}_{1,2} + \tilde{\pi}_2}\right). \end{aligned}$$

A.6. Conditions for Unicity of Steady State. For $\rho = 0$ the recursive optimization in (A.1) collapses to the standard Bellman equation of the linear quadratic Gaussian (LQG) problem with time-discounting. Theorem 3.4.1 (page 39) in [**?**] spells out the conditions for the existence and unicity of the steady state. For $\delta = 1$, if: i) the matrix $\mathbf{Q}$ is positive definite; ii) the matrix $\mathbf{R}$ is positive definite in $\{\mathbf{A}^m\}$, in that for some r $\sum_{m=0}^{r-1}(\mathbf{A}')^m\mathbf{R}(\mathbf{A})^m$ is positive definite; and iii) the matrix $\mathbf{J} = \mathbf{B}\mathbf{Q}^{-1}\mathbf{B}'$ is positive definite in $\{\mathbf{A}^m\}$, in that for some r $\sum_{m=0}^{r-1}(\mathbf{A}')^m\mathbf{J}(\mathbf{A})^m$ is positive definite, then the Riccati equation

$$\mathbf{\Pi} = \mathbf{R} + \mathbf{A}'\mathbf{\Pi}\mathbf{A} - \mathbf{A}'\mathbf{\Pi}\mathbf{B}(\mathbf{Q} + \mathbf{B}'\mathbf{\Pi}\mathbf{B})^{-1}\mathbf{B}'\mathbf{\Pi}\mathbf{A}\,,$$

possesses a unique semi-positive definite solution, $\mathbf{\Pi}$. Whittle discusses how to modify these conditions for the class of linear exponential quadratic Gaussian (LEQG) problems. In Theorem 9.2.1 (page 118) he states that if conditions i), ii) and iii), with $\mathbf{J}$ now equal to $\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}' - \rho\mathbf{N}$, alongside condition iv) that $\mathbf{J}$ being positive definite, hold, then the modified Riccati equation,

$$\begin{aligned} \mathbf{\Pi} &= \mathbf{R} + \mathbf{A}'\widetilde{\mathbf{\Pi}}\mathbf{A} - \mathbf{A}'\widetilde{\mathbf{\Pi}}\mathbf{B}(\mathbf{Q} + \mathbf{B}'\widetilde{\mathbf{\Pi}}\mathbf{B})^{-1}\mathbf{B}'\widetilde{\mathbf{\Pi}}\mathbf{A}\,, \quad \text{with} \qquad (A.6) \\ \widetilde{\mathbf{\Pi}} &= (\mathbf{\Pi}^{-1} - \rho\,\mathbf{N})^{-1}\,, \end{aligned}$$

possesses a unique semi-positive definite solution $\mathbf{\Pi}$. To adapt this result to the class of discounted LEQG (DLEQG) problems analyzed by [?] and [?], we employ a useful identity formerly exploited by [?] in the proof of Theorem 3.5.1 (pages 40-41). In particular, it is immediate to see that

$$\mathbf{z}'\mathbf{\Pi}\mathbf{z} \;=\; \max_{\boldsymbol{\mu}}(-2\boldsymbol{\mu}'\mathbf{z} \;-\; \boldsymbol{\mu}'\mathbf{\Pi}^{-1}\boldsymbol{\mu})\,.$$

In the LQG problem the matrix $\mathbf{\Pi}_t$ solves the Bellman equation (where one can appeal to the certainty equivalence principle and disregard the idiosyncratic shocks)

$$\mathbf{z}'\mathbf{\Pi}_t\mathbf{z} \;=\; \min_{u_t}\left[\, c(\mathbf{z}_t,u_t) \;+\; (\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t)'\mathbf{\Pi}_{t+1}(\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t)\,\right],$$

with $c(\mathbf{z}_t,u_t)=\mathbf{z}_t'\mathbf{R}\mathbf{z}_t+\mathbf{Q}u_t^2$. Considering the former identity we can write this as

$$\mathbf{z}'\mathbf{\Pi}_t\mathbf{z} \;=\; \max_{\boldsymbol{\mu}}\;\min_{u_t}\left[\, c(\mathbf{z}_t,u_t) \;-\; 2\,\boldsymbol{\mu}'(\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t) \;-\; \boldsymbol{\mu}'\mathbf{\Pi}_{t+1}^{-1}\boldsymbol{\mu}\,\right].$$

Because the argument in the brackets is convex in u_t and concave in $\boldsymbol{\mu}$ it admits a unique saddle point. This implies that one can invert the order of optimization. Via this transformation Whittle proves that the matrix $\mathbf{\Pi}_t$ also respects the following recursion

$$\mathbf{\Pi}_t \quad = \quad \mathbf{R} \;+\; \mathbf{A}'(\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}'+\mathbf{\Pi}_{t+1}^{-1})^{-1}\mathbf{A}\,.$$

This is an alternative formulation of the Riccati equation for the LQG problem. It roots out the salient elements which pin down the existence and unicity conditions of a steady state solution.

In the analysis of the LEQG problem Whittle shows that the matrix $\mathbf{\Pi}_t$ solves the following recursion

$$\mathbf{z}'\mathbf{\Pi}_t\mathbf{z} \;=\; \min_{u_t}\;\max_{\boldsymbol{\epsilon}_{t+1}}\left[\, c(\mathbf{z}_t,u_t) \;-\; (\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t+\boldsymbol{\epsilon}_{t+1})'\mathbf{\Pi}_{t+1}(\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t+\boldsymbol{\epsilon}_{t+1}) \;-\; \frac{1}{\rho}\,\boldsymbol{\epsilon}_{t+1}'\mathbf{N}^{-1}\boldsymbol{\epsilon}_{t+1}\,\right].$$

Maximizing the argument in the brackets with respect to $\boldsymbol{\epsilon}_{t+1}$ one finds that $\mathbf{\Pi}_t$ solves the modified Bellman equation

$$\mathbf{z}'\mathbf{\Pi}_t\mathbf{z} \;=\; \min_{u_t}\left[\, c(\mathbf{z}_t,u_t) \;+\; (\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t)'\widetilde{\mathbf{\Pi}}_{t+1}(\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t)\,\right],$$

where $\widetilde{\mathbf{\Pi}}_{t+1}=(\mathbf{\Pi}_{t+1}^{-1}-\rho\mathbf{N})^{-1}$. Then, applying Whittle's identity, one can verify that

$$\begin{aligned}\mathbf{\Pi}_t \quad &= \quad \mathbf{R} \;+\; \mathbf{A}'(\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}'+\widetilde{\mathbf{\Pi}}_{t+1}^{-1})^{-1}\mathbf{A} \\ &= \quad \mathbf{R} \;+\; \mathbf{A}'\left(\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}'+\mathbf{\Pi}_{t+1}^{-1} \;-\rho\,\mathbf{N}\right)^{-1}\mathbf{A}\,.\end{aligned}$$

This shows that in the LEQG problem the matrix $\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}'-\rho\,\mathbf{N}$ replaces the matrix $\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}'$ in defining the conditions for the existence and the unicity of a steady state solution. [?] proves that in the DLEQG problem the matrix $\mathbf{\Pi}_t$ solves a recursion very similar to that which applies to the LEQG formulation. In fact,

$$\mathbf{z}'\mathbf{\Pi}_t\mathbf{z} \;=\; \min_{u_t}\max_{\boldsymbol{\epsilon}_{t+1}}\left[c(\mathbf{z}_t,u_t) \;-\; \delta(\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t+\boldsymbol{\epsilon}_{t+1})'\mathbf{\Pi}_{t+1}(\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t+\boldsymbol{\epsilon}_{t+1}) \;-\; \frac{1}{\rho}\,\boldsymbol{\epsilon}'_{t+1}\mathbf{N}^{-1}\boldsymbol{\epsilon}_{t+1}\right].$$

As shown in the proof of Theorem 2 in [?], maximizing the argument in the brackets with respect to $\boldsymbol{\epsilon}_{t+1}$ one finds that $\mathbf{\Pi}_t$ solves the modified Bellman equation

$$\mathbf{z}'\mathbf{\Pi}_t\mathbf{z} \;=\; \min_{u_t}\left[\,c(\mathbf{z}_t,u_t) \;+\; (\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t)'\widetilde{\mathbf{\Pi}}_{t+1}(\mathbf{A}\mathbf{z}_t+\mathbf{B}u_t)\,\right],$$

where $\widetilde{\mathbf{\Pi}}_{t+1} = ((\delta\mathbf{\Pi}_{t+1})^{-1} - \rho\mathbf{N})^{-1}$. Then, applying Whittle's identity, one can show that

$$\begin{aligned}
\mathbf{\Pi}_t &= \mathbf{R} \;+\; \mathbf{A}'(\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}' + \widetilde{\mathbf{\Pi}}_{t+1}^{-1})^{-1}\mathbf{A} \\
&= \mathbf{R} \;+\; \mathbf{A}'\left(\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}' + \frac{1}{\delta}\,\mathbf{\Pi}_{t+1}^{-1} \;-\rho\,\mathbf{N}\right)^{-1}\mathbf{A} \\
&= \mathbf{R} \;+\; \sqrt{\delta}\mathbf{A}'\left((\sqrt{\delta}\mathbf{B})\mathbf{Q}^{-1}(\sqrt{\delta}\mathbf{B})' + \mathbf{\Pi}_{t+1}^{-1} \;-\delta\rho\,\mathbf{N}\right)^{-1}\sqrt{\delta}\mathbf{A}.
\end{aligned}$$

We conclude that in the DLEQG problem if: i) the matrix $\mathbf{Q}$ is positive definite; ii) the matrix $\mathbf{R}$ is positive definite in $\{(\sqrt{\delta}\mathbf{A})^t\}$, in that for some r $\sum_{m=0}^{r-1}(\sqrt{\delta}\mathbf{A}')^m\mathbf{R}(\sqrt{\delta}\mathbf{A})^m$ is positive definite; and iii) the matrix $\mathbf{J} = (\sqrt{\delta}\mathbf{B})\mathbf{Q}^{-1}(\sqrt{\delta}\mathbf{B})' - \delta\rho\mathbf{N}$ is semi-positive definite and positive definite in $\{(\sqrt{\delta}\mathbf{A})^m\}$, in that for some r $\sum_{m=0}^{r-1}(\sqrt{\delta}\mathbf{A}')^m\mathbf{J}(\sqrt{\delta}\mathbf{A})^m$ is positive definite, then the modified Riccati equation (A.6), with

$$\widetilde{\mathbf{\Pi}} \;=\; ((\delta\mathbf{\Pi})^{-1} \;-\; \rho\,\mathbf{N})^{-1},$$

possesses a unique semi-positive definite solution $\mathbf{\Pi}$. Given that in our formulation of the DLEQG problem $\mathbf{Q} = \alpha > 0$, condition i) is obviously satisfied. Then, consider $(\sqrt{\delta}\mathbf{A}')^m\mathbf{R}(\sqrt{\delta}\mathbf{A})^m$. We suppose that

$$(\sqrt{\delta}\mathbf{A}')^m\mathbf{R}(\sqrt{\delta}\mathbf{A})^m \;=\; \delta^m\,\beta\begin{pmatrix}\gamma^{2m} & D_{m-1}\gamma^m \\ D_{m-1}\gamma^m & D_{m-1}^2\end{pmatrix} \quad \text{where } D_m \;=\; D_{m-1} \,+\, \gamma^m \;\text{ and }\; D_1 = 1\,.$$

This conjecture is obviously true for $m = 1$. To check that this is correct for any other m consider that it implies that

$$
\begin{aligned}
(\sqrt{\delta}\mathbf{A}')^{m+1}\mathbf{R}(\sqrt{\delta}\mathbf{A})^{m+1} &= \delta^{m+1}\,\beta \begin{pmatrix} \gamma & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \gamma^{2m} & D_{m-1}\gamma^{m} \\ D_{m-1}\gamma^{m} & D_{m-1}^2 \end{pmatrix} \begin{pmatrix} \gamma & 1 \\ 0 & 1 \end{pmatrix} \\
&= \delta^{m+1}\,\beta \begin{pmatrix} \gamma^{2(m+1)} & \gamma^{m+1}(D_{m-1}+\gamma^{m}) \\ \gamma^{m+1}(D_{m-1}+\gamma^{m}) & (D_{m-1}+\gamma^{m})^2 \end{pmatrix} \\
&= \delta^{m+1}\,\beta \begin{pmatrix} \gamma^{2(m+1)} & D_{m}\gamma^{m+1} \\ D_{m}\gamma^{m+1} & D_{m}^2 \end{pmatrix},
\end{aligned}
$$

which is consistent with our initial conjecture. On the basis of this result we see that

$$
\mathbf{z}_0' \sum_{m=0}^{r-1} (\sqrt{\delta}\mathbf{A}')^{m}\mathbf{R}(\sqrt{\delta}\mathbf{A})^{m}\mathbf{z}_0 \;=\; \beta \sum_{m=0}^{r-1} \delta^{m}\,(\gamma^{m} p_0 \,+\, D_{m-1}e_0)^2 \,.
$$

Since for any choice of p_0 and e_0 there is at least a natural number m such that $(\gamma^m p_0 + D_{m-1}e_0) \neq 0$, we conclude that this value is strictly positive and hence that condition ii) is satisfied.

Now, consider $(\sqrt{\delta}\mathbf{A}')^{m}(\sqrt{\delta}\mathbf{B})\mathbf{Q}^{-1}(\sqrt{\delta}\mathbf{B})'(\sqrt{\delta}\mathbf{A})^{m}$. Suppose that

$$
(\sqrt{\delta}\mathbf{A}')^{m}(\sqrt{\delta}\mathbf{B})\mathbf{Q}^{-1}(\sqrt{\delta}\mathbf{B})'(\sqrt{\delta}\mathbf{A})^{m} \;=\; \delta^{m+1}\,\frac{1}{\alpha} \begin{pmatrix} \gamma^{2m} & S_{m-1}\gamma^{m} \\ S_{m-1}\gamma^{m} & S_{m-1}^2 \end{pmatrix} \quad \text{where } \; S_m \;=\; 1 + D_m \,.
$$

Once again, it is immediate to verify that this conjecture is true for $m = 1$. To check it is valid for any other m consider that it consistently entails that

$$
\begin{aligned}
(\sqrt{\delta}\mathbf{A}')^{m+1}(\sqrt{\delta}\mathbf{B})\mathbf{Q}^{-1}(\sqrt{\delta}\mathbf{B})'(\sqrt{\delta}\mathbf{A})^{m+1} &= \delta^{m+2}\,\frac{1}{\alpha} \begin{pmatrix} \gamma & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \gamma^{2m} & S_{m-1}\gamma^{m} \\ S_{m-1}\gamma^{m} & S_{m-1}^2 \end{pmatrix} \begin{pmatrix} \gamma & 1 \\ 0 & 1 \end{pmatrix} \\
&= \delta^{m+2}\,\frac{1}{\alpha} \begin{pmatrix} \gamma^{2(m+1)} & \gamma^{m+1}(S_{m-1}+\gamma^{m}) \\ \gamma^{m+1}(S_{m-1}+\gamma^{m}) & (S_{m-1}^2+\gamma^{m})^2 \end{pmatrix} \\
&= \delta^{m+2}\,\frac{1}{\alpha} \begin{pmatrix} \gamma^{2(m+1)} & S_{m}\gamma^{m+1} \\ S_{m}\gamma^{m+1} & S_{m}^2 \end{pmatrix}.
\end{aligned}
$$

Exploiting this result we see that

$$
\mathbf{z}_0' \sum_{m=0}^{r-1} (\sqrt{\delta}\mathbf{A}')^{m}(\sqrt{\delta}\mathbf{B})\mathbf{Q}^{-1}(\sqrt{\delta}\mathbf{B})'(\sqrt{\delta}\mathbf{A})^{m}\mathbf{z}_0 \;=\; \frac{1}{\alpha} \sum_{m=0}^{r-1} \delta^{m+1}\,(\gamma^{m} p_0 \,+\, S_{m-1}e_0)^2 \,.
$$

Since for any p_0 and e_0 there is at least a natural number m such that $(\gamma^m p_0 + S_{m-1}e_0) \neq 0$, we conclude that this value is strictly positive. Because $\mathbf{N}$ is finite, for ρ small enough we conjecture that

$\mathbf{J} \approx (\sqrt{\delta}\mathbf{B}')^m \mathbf{Q}^{-1}(\sqrt{\delta}\mathbf{B})$ is positive definite in $\{(\sqrt{\delta}\mathbf{A})^m\}$. A similar argument shows that for ρ small enough $\mathbf{J}$ is semi-positive definite and hence that condition iii) is also satisfied. In brief, we have checked that for ρ small, the steady state is unique.

A.7. The Coefficient ρ and the Relative Risk-aversion. Using results in [**?**], we have seen that the risk-enhancement coefficient is

$$\rho \;=\; 2\left(\frac{1}{\delta} - 1\right)(\chi - 1)\;.$$

This value is larger than zero if $\chi > 1 = 1/\theta$, i.e. if the coefficient of relative risk-aversion is larger than the inverse of the inter-temporal elasticity of substitution in Epstein and Zin's recursive preferences. In other words, a positive risk-enhancement coefficient is equivalent to the condition that the coefficient of relative risk-aversion is larger than the inverse of the inter-temporal elasticity of substitution. Interestingly, we can also write that

$$\chi \;=\; 1 \;+\; \frac{1}{2}\left(\frac{\delta}{1-\delta}\right)\rho\;.$$

This implies that in our base parametrization, given that $\delta = 0.97$, for ρ that ranges between 0 and 1, the coefficient of relative risk-aversion, χ, varies from 1 to 17.16. For $\delta = 0.95$ ($\delta = 0.99$), for ρ in the interval between 0 and 1, the coefficient of relative risk-aversion, χ, varies from 1 to 10.50 (1 to 50.50).

A.8. Unconditional Variance of Control Variable. Let $\mathbf{\Gamma} = \mathbf{A} + \mathbf{BK}$. Given $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{K} = (\kappa_p \;\, \kappa_e)$,

$$\mathbf{I}_2 \,-\, \mathbf{\Gamma} \;=\; \begin{pmatrix} (1-\gamma)-\kappa_p & -(1+\kappa_e) \\ -\kappa_p & -\kappa_e \end{pmatrix}.$$

Let $d = \det(\mathbf{I}_2 \,-\, \mathbf{\Gamma})$. This is $d = -(\kappa_p + (1-\gamma)\kappa_e)$. Then,

$$\mathbf{\Lambda} \;=\; (\mathbf{I}_2 \,-\, \mathbf{\Gamma})^{-1} \;=\; \frac{1}{d}\begin{pmatrix} -\kappa_e & 1+\kappa_e \\ \kappa_p & (1-\gamma)-\kappa_p \end{pmatrix}.$$

Therefore, $\mathbf{K}\,\mathbf{\Lambda} \;=\; \frac{1}{d}\,(0 \;\;\; (\kappa_p + (1-\gamma)\kappa_e) \;=\; (0 \;\;\; -1)$ and hence, given $\mathbf{N}$,

$$\begin{aligned} \mathrm{Var}\,[u_t] \;&=\; \mathbf{K}\,\mathbf{\Lambda}\,\mathbf{N}\,\mathbf{\Lambda}'\mathbf{K}' \\ &=\; (0 \;\;\; -1)\begin{pmatrix} \sigma_p^2+\sigma_e^2 & \sigma_e^2 \\ \sigma_e^2 & \sigma_e^2 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \end{pmatrix} \;=\; \sigma_e^2\,. \end{aligned}$$

A.9. The Coefficient ρ and Early Resolution of Uncertainty. In [**?**] it is noted that when the relative risk-aversion is greater than the inverse of the inter-temporal elasticity of substitution, i.e. for $\chi > 1/\theta$, the social planner's preferences favor early resolution of uncertainty *vis-a-vis* the standard case of expected utility. In fact, for $\chi = 1/\theta$ (or equivalently $\rho = 0$) Epstein and Zin's recursive preferences become linear, so that the utility function assumes the familiar time-separable form, while the value function solves the standard Bellman's equation from dynamic programming, $\boldsymbol{\mathcal{V}}_t = \min_{u_t} \{c_t + \delta E_t [\boldsymbol{\mathcal{V}}_{t+1}]\}$. In our specification for $\rho > 0$ we see that $\chi > 1$, where $1 = 1/\theta$. This means that, applying Kreps and Porteus's argument, we can affirm that the objective function in our recursive optimization induces earlier resolution of uncertainty *vis-a-vis* the case of expected utility.

A.10. Optimal Control in the Lagged Observation Scenario. From Theorem 4 and Lemma 7 in [**?**] we know that to find the optimal control when the social planner only observes a noisy signal on the state vector $\mathbf{z}_{t-1}$ in t, one needs to maximize with respect to $\mathbf{z}_t$ the sum

$$-(1/\rho)(\mathbf{z}_t - \hat{\mathbf{z}}_t)'\boldsymbol{\Omega}_t^{-1}(\mathbf{z}_t - \hat{\mathbf{z}}_t) \;+\; \mathbf{z}_t'\boldsymbol{\Pi}_t\mathbf{z}_t' \,,$$

where $\hat{z}_t$ is the maximum likelihood estimate of $\mathbf{z}_t$ at time t and $\boldsymbol{\Omega}_t$ is the corresponding conditional covariance matrix. When in t the social planner observes $\mathbf{z}_{t-1}$, $\mathbf{z}_t - \hat{\mathbf{z}}_t = \boldsymbol{\epsilon}_t$ and $\boldsymbol{\Omega}_t = \mathbf{N}$. Then, one needs to solve

$$\max_{\mathbf{z}_t} \left\{ -\frac{1}{\rho}\,(\mathbf{z}_t - \hat{\mathbf{z}}_t)'\mathbf{N}^{-1}(\mathbf{z}_t - \hat{\mathbf{z}}_t) + \mathbf{z}_t'\boldsymbol{\Pi}\mathbf{z}_t \right\} .$$

For $\mathbf{N}^{-1} - \rho\boldsymbol{\Pi}$ positive definite, this maximum is given for $\mathbf{z}_t = \check{\mathbf{z}}_t$, where

$$\check{\mathbf{z}}_t \;=\; (\mathbf{I} \,-\, \rho\mathbf{N}\,\boldsymbol{\Pi})^{-1}\hat{\mathbf{z}}_t \,.$$

Exploiting Theorem 4 in [**?**], it follows that the optimal control is then

$$u_t \;=\; \mathbf{K}\,\check{\mathbf{z}}_t \,,$$

where $\mathbf{K} = (\kappa_p \;\, \kappa_e)$, while κ_p and κ_e are given in Proposition 1.

A.11. Unconditional Variance of Control Variable in the Lagged Observation Scenario. In steady state, $\mathbf{z}_t = \mathbf{A}\mathbf{z}_{t-1} + \boldsymbol{\Psi}\hat{\mathbf{z}}_{t-1} + \boldsymbol{\epsilon}_t$, where $\boldsymbol{\Psi} = \mathbf{B}\mathbf{K}_I$. As the state vector is observed with a lag, $\hat{\mathbf{z}}_t = \mathbf{A}\mathbf{z}_{t-1} + \boldsymbol{\Psi}\hat{\mathbf{z}}_{t-1}$. Then, $\hat{\mathbf{z}}_t = \boldsymbol{\Phi}\mathbf{z}_{t-1}$, where $\boldsymbol{\Phi} = (\mathbf{I}_2 - \boldsymbol{\Psi})^{-1}\mathbf{A}$. Replacing this expression in that for $\mathbf{z}_t$ we find that $\mathbf{z}_t = \mathbf{A}\mathbf{z}_{t-1} + \boldsymbol{\Psi}\boldsymbol{\Phi}\,\mathbf{z}_{t-2} + \boldsymbol{\epsilon}_t$, which we can also write as $\mathbf{z}_t = (\mathbf{I}_2 - \mathbf{A}L - \boldsymbol{\Psi}\boldsymbol{\Phi}L^2)^{-1}\boldsymbol{\epsilon}_t$, so that $\mathrm{Var}[\mathbf{z}_t] = \boldsymbol{\Lambda}_I\mathbf{N}\boldsymbol{\Lambda}_I'$, where $\boldsymbol{\Lambda}_I = (\mathbf{I}_2 - \mathbf{A} - \boldsymbol{\Psi}\boldsymbol{\Phi})^{-1}$, $\boldsymbol{\Psi} = \mathbf{B}\mathbf{K}_I$ and $\boldsymbol{\Phi} = (\mathbf{I}_2 - \boldsymbol{\Psi})^{-1}\mathbf{A}$. In addition, as $\hat{\mathbf{z}}_t = \boldsymbol{\Phi}\mathbf{z}_{t-1}$ and $u_t = \mathbf{K}_I\hat{\mathbf{z}}_t$, $\mathrm{Var}[\hat{\mathbf{z}}_t] = \boldsymbol{\Phi}\boldsymbol{\Lambda}_I\mathbf{N}\boldsymbol{\Lambda}_I'\boldsymbol{\Phi}'$ and $\mathrm{Var}[u_t] = \mathbf{K}_I\boldsymbol{\Phi}\boldsymbol{\Lambda}_I\mathbf{N}\boldsymbol{\Lambda}_I'\boldsymbol{\Phi}'\mathbf{K}_I'$. Consider that

$$\mathbf{A} \,-\, \boldsymbol{\Psi}\,\boldsymbol{\Phi} \;=\; [\mathbf{I}_2 \,+\, \boldsymbol{\Psi}(\mathbf{I}_2 \,-\, \boldsymbol{\Psi})^{-1}]\mathbf{A} \,.$$

For any square matrix $\mathbf{M}$,

$$\mathbf{I} \,-\, \mathbf{M}\,(\mathbf{I} \,+\, \mathbf{M})^{-1} \;=\; \mathbf{I} + \mathbf{M} \,.$$

Taking $\mathbf{M} = -\boldsymbol{\Psi}$,

$$\mathbf{I}_2 \,+\, \boldsymbol{\Psi}(\mathbf{I}_2 \,-\, \boldsymbol{\Psi})^{-1} \;=\; (\mathbf{I}_2 \,-\, \boldsymbol{\Psi})^{-1}\,.$$

This implies that

$$\mathbf{I}_2 \,-\, \mathbf{A} - \boldsymbol{\Psi}\,\boldsymbol{\Phi} \;=\; \mathbf{I}_2 \,-\, (\mathbf{I}_2 - \boldsymbol{\Psi})^{-1}\mathbf{A} \;=\; \mathbf{I}_2 \,-\, \boldsymbol{\Phi}\,.$$

Hence, $\boldsymbol{\Lambda}_I = (\mathbf{I}_2 - \boldsymbol{\Phi})^{-1} = -\boldsymbol{\Phi}^{-1}(\mathbf{I}_2 - \boldsymbol{\Phi}^{-1})^{-1}$, where we have used the property that for $\mathbf{M}$ invertible, $(\mathbf{I} + \mathbf{M})^{-1} = \mathbf{M}^{-1}(\mathbf{I} + \mathbf{M}^{-1})^{-1}$. Then $\boldsymbol{\Phi}\boldsymbol{\Lambda}_I = -(\mathbf{I}_2 - \boldsymbol{\Phi}^{-1})^{-1}$, where $\boldsymbol{\Phi}^{-1} = \mathbf{A}^{-1}(\mathbf{I}_2 - \boldsymbol{\Psi})$. Let $\mathbf{K}_I = (\kappa_p^I \;\; \kappa_e^I)$. Given $\mathbf{B}$,

$$\mathbf{I}_2 \,-\, \boldsymbol{\Psi} \quad = \quad \begin{pmatrix} 1 - \kappa_p^I & -\kappa_e^I \\ -\kappa_p^I & 1 - \kappa_e^I \end{pmatrix}.$$

For

$$\mathbf{A}^{-1} \;=\; \frac{1}{\gamma}\begin{pmatrix} 1 & -1 \\ 0 & \gamma \end{pmatrix},$$

$$\mathbf{I}_2 \,-\, \mathbf{A}^{-1}(\mathbf{I}_2 \,-\, \boldsymbol{\Psi}) \;=\; \begin{pmatrix} 1 & \frac{1}{\gamma} \\ \kappa_p^I & \kappa_e^I \end{pmatrix} \quad \text{and} \quad (\mathbf{I}_2 \,-\, \mathbf{A}^{-1}(\mathbf{I}_2 \,-\, \boldsymbol{\Psi}))^{-1} \;=\; \frac{1}{\kappa_e^I - \frac{1}{\gamma}\kappa_p^I}\begin{pmatrix} \kappa_e^I & -\frac{1}{\gamma} \\ -\kappa_p^I & 1 \end{pmatrix}.$$

Finally,

$$\mathbf{K}_I\,\boldsymbol{\Phi}\,\boldsymbol{\Lambda}_I \quad = \quad -\,(\kappa_p^I \;\; \kappa_e^I)\,\frac{1}{\kappa_e^I - \frac{1}{\gamma}\kappa_p^I}\begin{pmatrix} \kappa_e^I & -\frac{1}{\gamma} \\ -\kappa_p^I & 1 \end{pmatrix} \;=\; (0 \quad -1)\,.$$

This implies that

$$\begin{aligned} \mathrm{Var}\,[u_t] \quad &= \quad \mathbf{K}_I\,\boldsymbol{\Phi}\,\boldsymbol{\Lambda}_I\,\mathbf{N}\,\boldsymbol{\Lambda}'\,\boldsymbol{\Phi}\,\mathbf{K}_I' \\ &= \quad (0 \quad -1)\begin{pmatrix} \sigma_p^2 + \sigma_e^2 & \sigma_e^2 \\ \sigma_e^2 & \sigma_e^2 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \end{pmatrix} \;=\; \sigma_e^2\,. \end{aligned}$$

B. Parametric Calibration

The parameters of the baseline configuration have been carefully calibrated as follows:

- The coefficient γ, which represents the rate of carbon decay, has been set equal to 0.9917, consistently with the value of the decay rate used in [**?**] and [**?**]. These authors borrow their choice from [**?**].

- The coefficient σ_p^2 is the variance of the shocks to the concentration level. We have set σ_p^2 equal to $0.0549 = 0.2343^2$, coherently with the value used by [?] for the standard deviation of the shocks to the pollution level, 0.2343. This value is equal to 0.11, the standard deviation for the concentration level measured in ppm estimated by the National Oceanic and Atmospheric Administration on the basis of Mauna Loa CO2 data for the period from 1959 to 2010, pre-multiplied by 2.13, the parameter converting ppm of atmospheric CO2 to Gigatonnes Carbon ($0.2343 = 2.13 \times 0.11$).

- The coefficient σ_e^2 is the variance of the shocks to the emission level. We have chosen a value of σ_e^2 equal to 0.01, so that $\sigma_e = 0.1$GtC. This choice implies that σ_e is lower than σ_p, coherently with the fact that e_t is lower than p_t in absolute terms, but also that σ_e, relatively to the mean values, is higher than σ_p, consistently with the assumption that society encounters difficulties in controlling emission flows. Differently from the stock of accumulated pollution, the extant literature does not consider uncertainty on emissions, so we have not found values that could be compared to σ_e^2. However similar results for the properties the optimal mitigation policy outlined in our analysis are obtained for alternative values of the parameter σ_e^2 as illustrated in the web Appendix.

- The coefficient α, that measures the direct cost of the abatement effort, has been set equal to 30. This implies a cost of reducing emissions by 1GtC equal to \$30 billions. In our numerical analysis this value for α brings about an optimal abatement effort ranging from -4.4GtC to -6.7GtC (depending on the choice of the discount factor δ and the risk-enhancement coefficient ρ). Given the strong convexity of the mitigation costs, such mitigation efforts correspond to an annual abatement cost ranging from \$580 to \$1390 billions, equal respectively to 0.8 and 1.9 percent of the gross world product (GWP) in 2012. These values seem both reasonable and coherent with the idea that mitigation effort does not necessarily impair economic activity. Notice that since there is no consensus on the costs of mitigation policies (see, for instance, [?]), we deem important to replicate our numerical analysis for different values of α. Crucially, in the web Appendix we show that the patterns of the abatement effort are qualitatively similar under different values of α.

- The coefficient β, that measures the economic cost of pollution, has been set equal to 0.01115. Given a deviation of the concentration level from pre-industrial values equal to 246GtC in 2012, such value for β implies an economic cost of pollution \$675 billions ($= 0.01115\$\text{billion}/\text{GtC}^2 \times (246\text{GtC})^2$). Our choice of β is equal to the coefficient $g/2$ of the quadratic environmental damage function of [?], where $g = 0.0223$ is the value used in their calibration. This value is derived in [?] (see their web Appendix) as they set $g = \frac{(20)(29185)}{(590)^2}10^{-2}(1.33)$, where 1.33 is the expected percentage reduction of Gross World Product (GWP) due to a doubling of CO2 atmospheric stocks over a period of ten years from their pre-industrial level, 590 is the pre-industrial atmospheric stock in GtC, and \$29,185 billions is the 1998 estimate of GWP (International Monetary Fund, 1999).

- The coefficient δ is the discount factor. We consider three values, $\delta = 0.95$, $\delta = 0.97$ and $\delta = 0.99$, in our numerical analysis. The mid value, which we use in our baseline calibration, coincides with the rate of time preferences adopted by [**?**].

- Finally, in our numerical analysis we consider values of ρ, the risk-enhancement coefficient, ranging from 0 to 1. The mid value, which we employ in our baseline calibration, corresponds to a coefficient of relative risk-aversion equal to about 9, as explained in Appendix A.7.